6-2015

MICROSATELLITE ANALYSIS OF POPULATION STRUCTURE IN THE SANTA ANA SPECKLED DACE (RHINICTHYS OSCULUS)

Stacey A. Nerkowski
California State University - San Bernardino, nerkowss@csusb.edu

MICROSATELLITE ANALYSIS OF POPULATION STRUCTURE IN THE SANTA ANA SPECKLED DACE (*RHINICTHYS OSCULUS*)

A Thesis

Presented to the

Faculty of

California State University,

San Bernardino

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

in

Biology

---

by

Stacey Ann Nerkowski

June 2015

MICROSATELLITE ANALYSIS OF POPULATION STRUCTURE IN THE SANTA ANA SPECKLED DACE (*RHINICHTHYS OSCULUS*)

A Thesis

Presented to the

Faculty of

California State University,

San Bernardino

---

by

Stacey Ann Nerkowski

June 2015

Approved by:

Dr. Anthony Metcalf, Committee Chair, Biology

Dr. James Ferrari, Committee Member

Dr. Dave Polcyn, Committee Member

ABSTRACT

Rhinichthys osculus, the Speckled Dace, is one of the most ubiquitous fish in western North America. Within the Southern California region, the local taxon is known as the Santa Ana Speckled Dace. The purpose of this study was to characterize and identify polymorphic microsatellite markers for *R. osculus* in which twenty-three were identified through Illumina pair-end sequencing. Seven of these loci were then used to examine the patterns of genetic variation and population structure that occurred within and among the watersheds in the Southern California. The study also examined the regional relationships among Southern California, Central California and Owen's River Valley. Analysis of the microsatellite data revealed highly significant moderate levels of population structure exist within the Southern California region (R_{ST}=0.160, p=0.001). This structure is best explained by watershed as well as isolation by distance (R^2=.2286, p=0.010). Highly significant geographic structure also exists among the geographic regions of Southern California, Central Coast, and Owen's River Valley regions ($R_{ST=}$ 0.600, p-value=0.001) that are congruent with the regional differentiation elucidated by mtDNA sequence data. In both cases, the degree of population differentiation was correlated with isolation by distance. Utilizing this information we were able gain a better understanding of the evolutionary relationships among the Southern California populations of Santa Ana Speckled Dace. Within the Santa Ana Speckled Dace populations we examined four models to explain the geographic structure: watershed, mountain range,

tributary, and isolation by distance. While all were significant, the tributary model exhibited the higher level of population structure ($R_{ST=}$ 0.160, p-value=0.001) and a significant correlation was exhibited between geographic distance and population structure, suggesting isolation by distance may be playing a role. The results of the microsatellite analysis are congruent with an earlier broad scale analysis of mtDNA sequence data that suggests the Central California and the Owens Valley populations diverged from each other prior to the divergence of the Santa Ana Speckled Dace populations from the Colorado Basin populations, and that the Central Coast populations were not established as a result of a migration event from the Southern California populations, as was previously hypothesized. Primarily due to human activity, Santa Ana Speckled Dace habitat has become highly fragmented resulting in some populations becoming extirpated. We hope this study will guide the strategies for the conservation of the remaining populations of Santa Ana Speckled Dace and watershed management in Southern California.

.

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank the individuals and organizations that have supported and funded my thesis project which would not have been possible without them.

I would first like to thank my thesis advisor, Tony Metcalf, for the funding and support of my project. Many aspects of this project would not have been made possible without his guidance and expertise. I would also like to thank my committee members, Jim Ferrari and Dave Polcyn, for their contributions and patience throughout the various phases of my project.

As for specimen acquisition, I would like to thank the members of the United States Forest Service and California Department of Fish and Wildlife. Without the specimen collection from various locations throughout California, this study would not have been possible.

In addition, I would like to thank my lab mates, past and current, for their continued support and understanding through a long and grueling process. Jay and Pia VanMeter provided me with my first tutorials into the molecular workings of our lab. I would like to thank Nguyen Tran for his assistance in the DNA extractions of all of my samples and Joe Riley for his assistance in specimen acquisition and molecular protocols.

I owe much of my still remaining sanity to my lab mates and close friends, Liane Greaver, Suzy Neal and Stephanie Arnold. Our “coffee breaks” and outings provided the much needed moral support and time away from the lab.

I would finally like to thank my parents, Kim and Jerry Nerkowski, for assisting and supporting in this endeavor. I will be forever grateful for the love and support that you provided during this chapter of my life. I would also like to thank my aunt, Lisa Nerkowski, for her continuous support and understanding of my higher educational goals.

Lastly, I would like to thank the following organizations for their financial support of my project: United States Forest Service San Bernardino Branch; California State University, San Bernardino's Biology Department; California State University, San Bernardino ASI, Inc.; California State University, San Bernardino Office of Student Research; California State University, San Bernardino Water Resource Institute.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

Overview

The degree of genetic differentiation and population substructure has been well examined in studies of vertebrate animal populations. The degrees of sequence divergence of particular genes have provided researchers with a portrait of the evolutionary relationships between species and their times since divergence from their common ancestors. The geographical distribution of a particular species can be associated with unique patterns of genetic differentiation among subpopulations. California encompasses some of the most geographically complicated patterns of genetic diversity on earth, and the California Floristic Province (including Southern California) is considered one of the world's 25 most biologically rich and endangered terrestrial ecoregions (Meyes, Mittermeier, Mittermeier, Fonseca, & Kent, 2000). Widespread species in the Southern California region have also shown pronounced fine-scale genetic differences among populations, possibly resulting from a highly varied and fragmented landscape ((Harrison S. , 1999; Brodie III & Brodie Jr., 1990; Brown, Leebens-Mack, Thompson, Pellmyr, & Harrison, 1997; Maldonado, Davila, Stewart, Geffen, & Wayne, 1995; Metcalf, Nunney, & Hyman, 2001; Phillipsen & Metcalf, 2009; Vandergast, Bohonak, Weissman, & Fisher, 2007; Rodriguez-Robles, Denardo, & Staub, 1999).

Stream dwelling vertebrates, especially fish, provide an excellent model for examining such population substructure and genetic differentiation. Gene flow between tributaries and watersheds normally only occurs during times of flooding, allowing for these populations to be geographically isolated from one another for extended periods. Each population may be independently influenced by various evolutionary forces due to their isolation resulting in independent evolutionary lineages of genes to form.

Cyprinidae is one of the most diverse families of freshwater fish, and inhabits a variety of environments including lakes, ponds, creeks, tributaries and isolated springs across North America. Each of these environments contains different characteristics and histories. Due to these differences, the Cyprinidae family has proven to be a relevant model group to address the evolutionary effects of environmental and ecological factors (Simmons, Berendzen, & Mayden, 2003; Scott & Crossman, 1973)).

Rhinichthys osculus, the speckled dace, is one of the most ubiquitous freshwater fish in the western United States and occupies a variety of stream environments (Hubbs, Miller, & Hubbs, 1974). In the western United States, the speckled dace is the only native fish to be represented in the majority of the drainage systems (Miller R. R., 1958). Locally, the Santa Ana Speckled Dace is found in such creeks as Lytle Creek, Cajon Creek, Plunge Creek and Mill Creek (Figure 2). The speckled dace is a small, cyprinid fish that reaches a length of approximately 80mm (Moyle & Marchetti, 2006).

The local variation of the Speckled Dace is known as the Santa Ana Speckled Dace. The Santa Ana Speckled Dace's geographic range included the Santa Ana, San Gabriel and Los Angeles Watershed. Some populations have become extirpated, such as the Los Angeles River populations, as a result of anthropogenic effects such as the urbanization of the watershed and county controlled flood systems (Santa Ana River Watershed Project Authority, 2004). As the Santa Ana Speckled Dace's habitats become more and more fragmented, the populations are highly affected by climatic events such as fire and floods. As a result of their highly fragmented habitats and the extirpation of populations, the Santa Ana Speckled Dace was listed as a species of concern by the California Department of Fish and Wildlife (1995) and the United States Forest Service (1998). Due to the lack of peer reviewed descriptions of the Santa Ana Speckled Dace's taxonomic status including genetic data, they were never federally listed under the Endangered Species Act (Moyle & Marchetti, 2006).

Research is currently being performed in the lab of Dr. Anthony Metcalf at California State University, San Bernardino, on the *cytochrome b and d-loop* regions of mitochondrial DNA (mtDNA) of the Santa Ana Speckled Dace. In addition to the data acquired from the mtDNA sequences, this study proposes to analyze population level variation in this species using microsatellite loci. Microsatellite loci are 1-6 base tandem repeats found in nuclear DNA. They are useful for population genetic analysis because they are highly polymorphic, typically having 3 to 11 alleles with different numbers of repeats. The number of

times it repeats can vary between individuals, populations and even between species allowing it to be utilized as a fine scale marker (Hardy, Charbonnel, Fréville, & Heuertz, 2003).

We have isolated and characterized 23 polymorphic microsatellite loci for the Santa Ana Speckled Dace (Nunziata, Lance, Jones, Nerkowski, & Metcalf, 2013). The objective of this study is to utilize seven of the polymorphic microsatellite loci on samples of *R*. *osculus* representing the Southern California, Central Coast of California and Owens Valley populations. Utilizing the microsatellite data, we will examine gene flow and historic patterns of interbreeding among creeks to provide a genetic baseline for the tributaries within the Southern California region. In addition, comparing the Santa Ana Speckled Dace populations to those of other locations within California may provide insight into relationships and population structure among regions. With these data, we hope to gain a better understanding of the evolutionary histories and developments that have occurred among the local populations of *R. osculus*.

Phylogeography

Phylogeography is an interdisciplinary field of study that examines the processes leading to the geographic distribution and evolutionary relationships of populations and species, especially those of closely related species. Phylogeography expanded upon the concepts of historical biogeography, examining the geographical distributions of species utilizing climatic, geological, and other environmental forces. Phylogeography requires extensive input from a

variety of disciplines but especially those of molecular genetics/phylogenetics and population genetics. By analyzing the geographic distributions of current populations, Avise et al. (1987) stated that “one can begin to have a better understanding of the evolutionary footprints that may remain from ancestral populations that lead to the formation of the current population.”

A species’ geographic distribution typically results from geological and landscape barriers; these are barriers that result from physical topography changes such as the formation of mountain ranges, the creation of a river or lake, or the presence of a desert or elevation gradient. Each of these barriers can limit the migration and expansion of a population but can also cut off a population from its original source population. Over time, isolated populations may have been subjected to varying environmental stressors resulting in the population evolving independently from the ancestral population. This would result in a genetically distinct population which can be used in the mapping of the genetic and geographic distribution (Avise J. C., 2000).

How did the extant species of today arise from ancestral populations? What pressures or circumstances arose that caused a population to become geographically isolated, thereby creating the right conditions for divergence and eventually reproductive isolation? Natural selection plays a crucial role in the formation of new species but it can also contribute to population-level structure observed in geographical distributions (Avise, 1994; Avise, 2000).

Two competing schools of thought that account for the origins of spatially disjunct populations are vicariance and dispersal biogeography. Vicariance distribution results in emergence of new populations from the ancestral population due to environmental forces. Little to no gene flow exists between the ancestral population and the newly immerged populations, resulting in further divergence. A more varied historical relationship could occur through the dispersalist approach. This approach relies on the limited gene flow that would occur between emerging populations as a result of the modifications that have arisen since divergence from the ancestral population (Avise, 2000).

Early classification and description of species was typically based on a systematic approach where behavioral and morphological characteristics were used. Furthermore, subspecies classification was based in part on the geographical locations in which specimens were collected. Today, species and subspecies classification further describes the organisms by comparing and contrasting gene flow at the molecular level (Avise, 2000).

Migration (gene flow) contributes to the genetic variation of a population by allowing genes from one population to be brought into another population. It acts as a homogenizing force of evolution that tends to keep subpopulations of a species more similar genetically, resulting in little population substructure. Populations that are historically isolated, with little to no gene flow, typically have greater divergence in selectively neutral and non-neutral genes. Migration will tend to counteract the effects of genetic drift and selection. Migration can affect

genetic drift, but its effect is dependent on the relative strength of genetic drift (population size) and the relative strength of migration (migration rate). As heterozygosity increases within a population, genetic differentiation decreases, creating a more homogenized population (Hartl & Clark, 2006). When small populations colonize new areas, rare alleles are likely to be lost and heterozygosity is expected to decrease as generations progress. The new populations diverge genetically from the ancestral population and from each other.

When a new population is established by a few individuals colonizing an unoccupied area, initial genetic changes are likely to occur solely as a result of random processes. Rare alleles are likely to be lost in the founding event, and heterozygosity will be reduced in subsequent generations (Hedrick, 1998; Nei, 1987). The level of genetic differentiation in a population can provide researchers with an estimate of the level of migration and isolation. Populations that have little to no gene flow through migration should show high levels of genetic differentiation. As a result of genetic differentiation and migration limitations, a population's genetic structure can be established and analyzed.

Population Genetics

Population genetics examines genetic principles such as Mendelian genetics that affect the entire population of organisms. Population genetic studies examine population genetic parameters including genetic structure,

gene flow and migration, population stability and demographic history, In addition they can be used to quantify the effects of habitat fragmentation to guide conservation efforts.

The allele frequencies within a population can be affected by five evolutionary forces; non-random mating, natural selection, mutation, genetic drift and migration. Each of these forces can place various pressures on a population but the effects on allele frequency are greatest when a population is isolated (Hartl & Clark, 2006). Freshwater vertebrate species are particularly characterized by geographic isolation where they tend to have fragmented habitats due to varying connectivity of drainage and watershed systems, mountain ranges and anthropogenic affects (Phillipsen & Metcalf, 2009). Through the analysis of molecular markers, one can analyze and describe the phylogeography and population genetic parameters of a species for both current and ancient isolation events.

Hardy-Weinberg Equilibrium Model

The Hardy-Weinberg Model examines the expected genotypic and allelic frequencies that occur in a population where random mating occurs. The Hardy-Weinberg Model is a reference model that allows for no evolutionary forces including no migration, no genetic drift with an infinite population size, no mutation, no selection and where mating is random. Heterozygosity can maintain the frequency of rare recessive alleles within a population. Heterozygosity can range in value from 0 to almost 1, where increasing values

represent an increase in heterozygosity frequency within a population. Higher than expected levels of heterozygosity, typically indicate high levels of immigration and gene flow (Hartl & Clark, 2006).

Genetic Structure: Fixation Index (F_{ST})

To examine and quantify the inbreeding effect of population substructure, Sewall Wright defined the fixation index. The fixation index, F_{ST}, allows the user to examine genetic differentiation of subpopulations within a total population. The F_{ST}, can range from 0 to 1, where an increasing value represents greater genetic structure (Wright, 1931). Analogs of F_{ST} have been developed to further analyze population structure with use of various molecular markers such a ϕ_{ST} (Excoffier, Smouse, & Quattro, 1992) and R_{ST} (Balloux & Goudet, 2002).

Gene Flow and Migration

As discussed above, gene flow contributes to the genetic variation of a population by allowing genes from one population to be brought into another population. Migration is the exchange/movement of genes (individuals) from one population to another. In populations that receive no immigration, the presence of a small number of individuals able to reproduce will limit the genetic variation that is exhibited within the population over generations. Wright (1931) introduced effective population size as the ideal population size that would undergo genetic drift at the same rate as exhibited by the actual population. In the case of mtDNA and N_e, genetic drift will have a four-fold greater effect on mtDNA than that of nuclear DNA. N_e can be calculated by NF/2 where NF is equal to the number of

breeding females in the population. Since mtDNA is haploid, it contains a single set of genes from the mother instead of the two paternal and two maternal gene copies received in nuclear DNA (Hartl & Clark, 2006). The effective population size that can be estimated from mtDNA is ¼ of that from nuclear autosomal sequences; this could result in an increased lineage assorting rate and allele extinction rate (Hartl & Clark, 2006).

Mitochondrial DNA

During the 1970's, population genetics began analyzing evolutionary relationships with the use of mitochondrial DNA (mtDNA) markers. Animal mtDNA is maternally inherited and has a non-recombining mode of inheritance (Dawid & Blackler, 1972), a rapid rate of evolution (an average rate of 0.02 substitutions per base pair per million years) (Brown, Matthew George, & Wilson, 1979) and extensive intraspecific polymorphism (Dawid & Blackler, 1972; Moritz, Dowling, & Brown, 1987). mtDNA quickly became the molecular marker of choice when analyzing phylogenetic relationships and in developing phylogeography.

In addition, mtDNA was useful in evaluating gene flow and geographic isolation because the effective population size (Ne) of mtDNA is ¼ of that of nuclear autosomal sequences, resulting in more rapid divergence of populations due to genetic drift (Hartl & Clark, 2006; Avise, 1994). Researchers could utilize this information to examine the degree of genetic drift and geographic variation

that lead to the divergence of the organism from its ancestral predecessors. The mitochondrion is capable of self-replication due to the possession of its own genome (Dawid & Blackler, 1972).

The mitochondrial genome is highly compact such that no introns or long intergenic stretches of non-coding sequences exist, compared to what is typically found in the nuclear genome. Animal mtDNA is uniform in size across most taxa, being approximately 16kb in length and having a very stable gene arrangement in vertebrates (Brown, Matthew George, & Wilson, 1979; Dawid & Blackler, 1972). Due to the small size of the mitochondrial genome, replication occurs at a much faster rate than would normally occur in the nuclear genome (Brown, Matthew George, & Wilson, 1979). Since mitochondrial DNA is maternally transmitted, the recombination of genomic haplotypes is almost non-existent (Brown, Matthew George, & Wilson, 1979; Dawid & Blackler, 1972; Moritz, Dowling, & Brown, 1987). Nuclear DNA would usually be subjected to the variation and shuffling that would occur from bi-parental genetic material during meiosis, whereas mtDNA lacks this occurrence. As a result, the variation that occurs within the genome is generally the result of a mutation (Brown, Matthew George, & Wilson, 1979; Dawid & Blackler, 1972; Moritz, Dowling, & Brown, 1987).

Over the past 30 years, mtDNA sequences have been used on numerous vertebrate taxa including mammals, reptiles and amphibians, and numerous invertebrate species. Typically, in most phylogeographic studies, mtDNA

haplotypes that are similar or identical will be spatially proximal to one another. Greater sequence divergence occurs in a genealogical split that typically results in the differentiation and isolation of a population in separate refugia (Avise, 2009). Usually, this isolation can align with historical terrain changes in the geological record and are often concordant with landscape level variation.

As advances were made in molecular methods, the investigation of geographic distribution and variation could be analyzed at the molecular level, shedding light on intraspecific phylogeographic structures. Typically, phylogeography has been used to analyze the evolutionary distinct population segments of a species, but it has also been used as a tool in analyzing historical gene flow between populations (Avise, Ball, & Arnold, 1988), estimating population sizes throughout a species evolutionary lineage (Avise, Neigel, & Arnold, 1984), as well as defining the evolutionary trajectory for a species (Sites Jr. & Crandall, 1997). In recent studies, mtDNA was used to analyze the phylogeography and population history of various mammals (Vila, et al., 1999; Paetkau, Calvert, Stiriling, & Strobeck, 1995), birds (Mila, Girman, Kimura, & Smith, 2000; Zink, 1996), amphibians (Phillipsen & Metcalf, 2009; Tan & Wake, 1995) and insects (Vandergast, Bohonak, Weissman, & Fisher, 2007).

Southern California Population Studies

Vicariant events throughout the geological history of California have played a crucial role in shaping the diversification and distributions of California's extant species. Cismontane California is divided by the Transverse Range and

the Peninsular Range. The Transverse Range is a group of mountains that exist in Southern California. This mountain range gets its name from its East-West orientation rather than the North-South orientation typically exhibited by mountain ranges throughout North America. The Peninsular Range runs predominantly north-south, and includes the Santa Ana Mountains, the San Jacinto Mountains and the Laguna Mountains. It has been proposed that the transverse range has played an important role in the speciation and genetic distribution of species across Southern California. "Transverse Range Break " has been attributed to distinct north/south and/or east/west lineage breakages of numerous species.

The use of mtDNA was utilized to examine lineage breaks within the Transverse Mountain range. Chatzimanolis & Caterino (2007) examined the phylogeographic structure and localized lineage breaks of the rove beetle, *Sepedophilus castaneus,* through the amplification of the *Cytochrome Oxidase I* (*CO1*) gene in mtDNA. Their results revealed significant genetic and geographic structure among the *S. castaneus* populations of central and Southern California. Haplotypes fell into four distinct clades which represented Southwestern Sierra Nevada, Saint Lucias and two across the transverse range. Upon further analysis of their data, researchers believe an ancient isolation existed between the Sierra Nevadan population and that of the Eastern Transverse Range (Chatzimanolis & Caterino, 2007).

Research utilizing mtDNA has also revealed the genetic differentiation between populations and regions for the Jerusalem cricket in Southern

California. Vandergast, Bohonak, Weissman, & Fisher(2007) examined the phylogeography and genetic differentiation among current and prehistoric populations of *Stenopelmatus mahogany*. mtDNA's cytochrome oxidase gene was sequenced and analyzed for each sample. High elevation haplotype surrounding the Los Angeles basin represented a single clade, and other regional populations represented corresponding clades. One haplotype was represented in both the San Bernardino and Tequesquite sampling sites, which provided evidence into the implication that a single long distance dispersal event led to the creation of the Tequesquite population (Vandergast, Bohonak, Weissman, & Fisher, 2007).

The Transverse Range has been known to can act as a geographic and isolating barrier for one species but not affect the distribution of a similar species. Burns et al. (2006) examined the phylogeography of the wrentit, *Chamaea fasciata,* across the California Floristic Province. Within mtDNA, they amplified and sequenced protein-coding portions of *ATPase6, ATPase8, COII, COIII* and *cytochrome b*. Researchers were able to identify 6 distinct clades that were restricted geographically. Upon further investigation, they suggested that the wrentit was isolated during the Pleistocene into the Southern refugia, and only recently has the wrentit undergone a range expansion. Researchers compared their data to similar data for the California Thrasher, which reflected a distinct lineage break between the Transverse Ranges, unlike that of the wrentit (Burns & Barhoum, 2006).

Most researchers examine the phylogeography and genetic distribution of terrestrial animals, yet limited studies have been performed on the phylogeography of freshwater vertebrates of the Southern California region. Southern California is composed of numerous drainage and watersheds which could have been influenced by the landscape and geological history of the area. Gene flow between watersheds and drainages is based on periods of connectivity typically resulting from pluvial events.

Phillipsen & Metcalf (2009) examined the patterns of genetic variation that existed among putative landscape barriers across the range of the California Tree Frog, *Pseudacris cadaverina*. Using mtDNA cytochrome b sequences, they found that high levels of genetic differentiation existed in the Southern California populations as revealed by an AMOVA. Gene flow was limited between regional populations due to watersheds, mountain ranges and allopatric fragmentation between coastal and desert populations. Strong support for a Transverse Range Break among *P*. *cadaverina* was exhibited by their data (Phillipsen & Metcalf, 2009).

The majority of phylogeography studies conducted on the inhabitants of the Southern California region have utilized mtDNA as the molecular marker, and very little has been done using only nuclear DNA (nDNA) or a combination of nDNA and mtDNA. Spinks, Thomson, & Shaffer (2010) examined the phylogeography of the western pond turtle, *Emys marmorata*. Researchers utilized the ND4 site of mtDNA and five nuclear loci to map the phylogeography

within the California region. They utilized five nuclear loci that have relatively low intraspecific variation among populations, and compared them to the ND4 phylogeography trees. Their data suggested a north/south split among populations with an integration of the two Central Coast populations; this was further supported by the evidence of gene flow from the northern populations and the San Joaquin populations into the Central Coast region (Spinks, Thomson, & Shaffer, 2010). This study revealed the applications of nDNA in phylogeographic studies and enrichment as an applicable molecular marker. Appropriate, population level nDNA markers were found to be difficult to acquire in the early years of phylogeography, resulting in most of the published research utilizing mtDNA.

Mitochondrial DNA Disadvantages

Mitochondrial DNA continues to be essential in the field of phylogeography, yet it has many disadvantages that have lead researchers to the identification and development of nuclear DNA markers (Avise, et al., 1987; Avise, 1991; Harrison R. , 1989; Moritz, Dowling, & Brown, 1987; Simon, 1991). In animals, mtDNA pseudogenes, have been observed in the nuclear genome, which is highly undesirable in population genetic studies. Pseudogenes are sequences of DNA that exhibit gene-like qualities but are non-functional due to their lack of the ability to code for proteins. Corrections have been established to treat samples that may have mtDNA pseudogenes but these treatments have

shown limited effectiveness (Bensasson, Zhang, Hart, & Hewitt, 2001; Zhang & Hewitt, 1996).

An additional shortcoming of mitochondrial DNA is that it can be highly conserved between individuals of the same species and even those of the same genus. As a result, its uses in population genetic studies can be greatly limited among some animal taxa, such as those belonging to the grasshopper and locust order (Bensasson, Zhang, & Hewitt, 2000). Population structure would be difficult to establish when the mitochondrial genome is nearly identical in every individual of the population or species. The lack of recombination in the mitochondrial genome provides a limited and biased view of the evolutionary lineage of an organism, which could be very different when compared to the overall population or species lineage. In addition, due to the maternally biased evolutionary perspective, population structure could easily be affected by sex-biased dispersal (Bensasson D. , Zhang, Hart, & Hewitt, 2001; Zhang & Hewitt, 1996). The maternal inheritance only enables researchers to trace the maternal lineage through phylogenies and evolutionary relationships with other species/taxa (Forbes & Boyd, 1996; Schlötterer & Harr, 2000).

Microsatellites

nDNA is comprised of multiple genetic structures known as chromosomes. nDNA is bi-parentally inherited and in vertebrates encodes for the vast majority of genome when compared to mtDNA. It can be utilized as a molecular marker

when examining population genetics and phylogeography through direct assessment of the sequence variation. Numerous nDNA markers have been identified including restriction fragment length polymorphisms (RFLPs), variable number tandem repeats (VNTRs), and multi-locus and single-locus microsatellites. The rate of synonymous substitution for nDNA is 10^{-8} to 10^{-9} mutations per generation, making them effectively neutral in the presence of natural selection (Kumar & Subramanian, 2001).

The 1980's gave rise to the molecular revolution which provided researchers new powerful techniques and technologies based on the perception that polymorphisms could be used a molecular markers. A variety of molecular markers were identified on both mitochondrial DNA (mtDNA) as well on nuclear DNA (nDNA) which were utilized in population genetics and conservation biology. Due to the maternal bias in mtDNA, a new fine-scale molecular marker needed to be identified for use in phylogenetic relationships within species as well as interspecies relationships. Microsatellites became an important source of polymorphic genetic markers for the construction of linkage maps, parentage testing, genetic population structure and assessing evolutionary relationships.

Microsatellites are tandem repeats of 1 to 6 base pairs in nDNA scattered across chromosomes. In fish species, microsatellite arise approximately once every 10kbp (O'Connell & Wright, 1997). The lengths of the sequences are often di-, tri-, or tetra- nucleotide repeats which can be repeated up to approximately 100 times at a locus (Epifanio, Johnson, & Kassler, 2003). The repeat number

found at each locus can vary between an individual's alleles but also within the population and/or between species. Since nDNA is bi-parentally inherited (one copy is maternally inherited while the other is paternally inherited), an individual can vary in the number of repeat sequences for each allele, thus can either be homozygous or heterozygous at a specific locus. On this basis, we can use a variety of microsatellite loci to specify fingerprints which are used to identify specific individuals, as well as to employ population genetic analysis directly via heterozygosity (Al-Rabab'ah & Williams, 2002; Garcia-Moreno, Matocq, Roy, Geffen, & Wayne, 1996; Schlötterer & Harr, 2000).

Microsatellite mutation rates in repeat number range from 10^{-6} to 10^{-2} mutations per generation and are significantly higher than nuclear base substitution rates (O'Connell & Wright, 1997). The proposed mutation mechanism has been called "DNA (replication) slippage". It is assumed that during DNA replication the nascent and template strand realign out of normal alignment. If DNA synthesis continues on this molecule, the repeat number of the microsatellite is altered (Schlötterer & Harr, 2000).

Microsatellite markers are ideal for conservation biology due to their ability to be used to detect bi-parentally inherited polymorphisms as well as recent and past species admixture. Genetic admixture occurs when individuals from two or more previously geographically isolated populations begin interbreeding. Admixture results in the introduction of new genetic lineages into a population. The mutation rates and variability within microsatellite loci play important roles in

determining what type of analysis and research can be done with certain loci (Estoup & Angers, 1998; O'Connell & Wright, 1997)

Microsatellite loci are typically conserved in related taxa, but the repeat range will vary. Researchers can typically use microsatellites identified in a species and use them on related taxa in which microsatellites have not been characterized. Turner, Dowling, Broughton, & Gold (2004) used microsatellite loci that were isolated from one cyprinid fish, the common shiner (*Luxilus cornutus*) and attempted cross-species amplification in four other cyprinid species. Of the eight loci used in the experiment, five of the microsatellite loci produced "well-resolved, polymorphic and scorable products (Turner, Dowling, Broughton, & Gold, 2004)." Girard & Angers (2006) performed a similar experiment among several *Leuciscinae* species where they attempted cross-species amplification of 11 microsatellite loci isolated from *Rhinichthys cataractae* (Longnose Dace). Between four to ten loci were successfully amplified in five of the closely related species. Both of these cases provided microsatellite loci amplification across species that did not previously have microsatellite loci identified (Girard & Angers, 2006).

The advantages to the use of microsatellites have led to the gradual replacement of allozyme and mtDNA markers in population level and phylogeographic studies. Microsatellites are easy to amplify, due to their short fragment lengths, and relatively easy to isolate. In addition, very small amounts of tissue are required for the isolation of microsatellites. This allows for non-

lethal sample acquisition which extremely advantageous in populations that are endangered or have small population sizes (McConnell, O'Reilly, Hamilton, Wright, & Bentzen, 1995)

Two models that are appropriate for analyzing microsatellite data are the stepwise mutation model (SMM) and the two-phase model (TPM). The SMM suggests that in the evolution of microsatellites, they can only gain or lose a single repeat. This suggests that those individuals that differ by a single repeat are more closely related to one another than those that differ in larger repeat numbers (Jarne & Lagoda, 1996). The TPM incorporates the mutational processes of the SMM but allows for multiple mutations to occur instead of a single repeat number mutation (Di Rienzo, et al., 1994).

Genetic survey data can be collected and analyzed to infer the demography of a population (Avise, 1994). Population expansion and decline can be evaluated with microsatellite data by examining the level of heterozygosity (repeat numbers) that exists at each locus within a population. Based on the assumptions of mutation models, populations that are in decline reduce allelic diversity faster than heterozygosity or gene diversity (Nei, Maruyama, & Chakraborty, 1975). When the observed diversity in repeat number is greater than the expected diversity, a population may be declining in size. On the other hand, when allelic diversity is in deficit compared to the expected diversity, the population may currently be in a time of expansion (Shi, Kerdelhue, & Ye, 2012). Spatial Analysis of Molecular Variance (SAMOVA) can

then be utilized define groups of populations that are geographically homogeneous and maximally differentiated from each other, and the results can identify genetic barriers between these groups (Dupanloup, Schneider, & Excoffier, 2002).

Microsatellites can also have disadvantages, with the greatest being null alleles. Null alleles are non-amplifying alleles usually attributed to point mutations occurring at a primer annealing site. Microsatellite loci have been found to sometimes include a null allele. If a null allele is present and not accounted for it will create a false homozygote reading when scoring the genotypes. This mis-scoring will lead to a deficiency in the level of heterozygosity observed in the population. Null alleles can also cause erroneous results in the elimination of putative parents or the degree of relatedness between individuals in a kinship analysis (Ardren, Borer, Thrower, Joyce, & Kapuscinski, 1999; McConnell, O'Reilly, Hamilton, Wright, & Bentzen, 1995; O'Reilly & Wright, 1995). An additional disadvantage of microsatellites arises through stutter or shadow bands. During the amplification of a microsatellite loci via PCR, a ladder of bands 1-2 bps apart can be generated on the polyacrylamide gel; these are known as stutter or shadow bands. The bands typically result from slipped-strand mis-pairing that could occur during PCR. These stutter bands can result in mis-scoring of the alleles, yet O'Reilly & Wright (1995) observed that little to no stuttering occurs in tri- and tetra nucleotide repeat sequences.

Microsatellite Studies in Marine and Freshwater Vertebrates

Due to the high mutation rates exhibited by microsatellites, mtDNA is also needed to fully examine the phylogeography of many vertebrate species. mtDNA analysis contributes to the detection of older lineages within a species and can identify the relationships between species through a common ancestor. Microsatellites can provide insight into the population structure and gene flow that currently exists in a population, or changes that have occurred more recently in evolutionary time. Utilizing both of these molecular markers, researchers can shed much light on the evolution of particular species.

In fisheries science and conservation, allozyme and mitochondrial DNA genetic tags were used instead of co-dominant nuclear marker to differentiate and monitor fish populations. However, the problem that arose with these methods were that many required lethal sampling, and low genetic variability was often observed. Researchers in fisheries science have turned to microsatellite markers to examine fish populations. Borer, Miller, & Kapuscinski (1999) used microsatellite loci to examine variability among 186 individuals of walleye (*Stizostedion vitreum*) sampled from northern Minnesota. Based on their data, they were able to determine that there was very little relatedness among the individuals within and between the sampled populations and that within populations, high levels of genetic variation existed (Borer, Miller, & Kapuscinski, 1999).

Ward, Bowers, Hensley, Mobley, & Belouski (2007) examined microsatellite loci in the spotted sea trout (*Cynoscion nebulosus*) that were collected from nine different bays around Texas. Statistically significant differences were observed among the bays in which the upper Texas coast bays clustered together and the lower Texas coast bays clustered together. These findings disagreed with previous allozyme and mtDNA studies. Since allozyme and mtDNA are not fine-scale markers, they may not have been able to detect the different selection pressures currently in place against the populations (Ward, Bowers, Hensley, Mobley, & Belouski, 2007).

Microsatellites have also been used to examine the levels of genetic diversity that exist among populations in order to implement conservation policies. Turner, Dowling, Broughton, & Gold (2004) identified microsatellite loci from the common shiner, *Luxilus cornutus*, which were successfully amplified across four species among the subfamily *Leusicinae*. The microsatellite loci in this study revealed its utility in population and conservation genetics applications for numerous cyprinid fish species (Turner, Dowling, Broughton, & Gold, 2004).

Species diversity and geographic distribution were examined by Fernandes-Matioli, Matioli, & Almeida-Toledo (2000) using microsatellites in the genus *Gymnotus*, a monophyletic group of electrogenic freshwater fish in the Amazon basin. Researchers used microsatellites as flanking primer pairs to see what sequences were amplified in four different *Gymnotus* species. One of their primers amplified species-specific products in which researchers were able to

determine geographic distribution of *G. sylvius,* which was used to further describe the species (Fernandes-Matioli, Matioli, & Almeida-Toledo, 2000).

The California Floristic Province is a region with great biodiversity among terrestrial and aquatic plants and animals. Aguilar & Jones (2009) examined two native California minnows*, Lavina symmetricusI* (roach) and *Lavina exilicauda* (hitch), for nuclear (8 microsatellites) and mitochondrial (control region and ND2) diversification. mtDNA analysis revealed two distinct, highly divergent clades for *L. symmetricus,* representing the Gualala and Pit Rivers. It was estimated that these two clades diverged from an ancestral population of *Lavina* approximately 3-6 million years ago. The problem that arose was that there were no distinct haplotypes that could distinguish the two different species within the Sacramento and San Joaquin region. Upon analysis of the microsatellite data, the subspecies previously described for each species were identified as genetically distinct units and conservation management could then be implemented to manage them as such units (Aguilar & Jones, 2009).

Further investigation needs to be performed on freshwater vertebrate phylogeography for the Southern California region. For this reason, *Rhinichthys osculus*, the speckled dace, would prove to be a valuable species in further examining the divergence patterns that exist among the watershed populations of Southern California using both mtDNA and nDNA. In addition, gene flow can be examined for the drainage systems that currently house populations of *R. osculus*. With this information we can provide further insight into the

phylogeography and geographic distribution of our local populations for conservation and reintroduction use.

Phylogeography of the Speckled Dace

Rhinichthys osculus, the speckled dace, is considered to be one of the most ubiquitous freshwater fish in the Western United States and occupies a variety of environments (Hubbs, Miller, & Hubbs, 1974) from the Columbia Basin in Canada to the Sonora Basin of Mexico (Lee, et al., 1980). In the Western United States, the speckled dace is the only native fish to be represented in the majority of the major drainage systems (Miller R. R., 1958). *Rhinichthys osculus* belongs to the Cyprinidae, one of the most diverse families of freshwater fish. The speckled dace ranges in size from 5-9 cm in length. They have morphological adaptations such as streamlined bodies and large falcate fins that allow them to swim in swift currents but also maintain their position when needed (Smith & Dowling, 2008). The current geographic distribution of the Speckled Dace in California is represented in Figure 1.

The species belonging to the Cyprinidae inhabit a variety of environments including lakes, ponds, creeks, tributaries and even isolated springs across North America. Species of the Cyprinidae family have proven to be relevant models to address environmental and ecological changes from an evolutionary perspective (Girard & Angers, 2006; Scott & Crossman, 1973; Simmons, Berendzen, & Mayden, 2003).

R. osculus was first identified as the western genus *Apocope* which researchers thought contained at least 12 different species. Upon further taxonomic analysis, *Apocope* became a sub-genus and the species are now considered a single-wide ranging species known as *Rhinichthys osculus* (Girard) (Jordan, Evermann, & Clark, 1930; Miller & Miller, 1948). Each geographical population adapted morphological characteristics that were better suited for that environment, giving rise to the vast morphological differentiation that exists between regional populations. Oakey, Douglas, & Douglas (2004) examined the molecular phylogeny of *R. osculus* and compared it to the hydrographic evolution of Western North America. Researchers mapped 112 restriction sites in mtDNA among 59 sampled populations. Upon construction of their phylogenetic trees, two large clades, representing the Colorado River clade and the Snake River clade, were predominant in their results. Within the major clades were smaller sub-basin clades that were created due to aridity, tectonics and elevation (Oakey, Douglas, & Douglas, 2004).

Smith & Dowling (2008) were able to determine that the estimated time of divergence for *R. osculus* from an ancestral fish species occurred approximately 6.3 million years ago (mya) in the Colorado River Basin. Based on mtDNA sequences of the cytochrome b region and ND4 site, they projected that *R. osculus* diverged from the Colorado River basin and into connecting watersheds (including the Los Angeles Basin) approximately 1.9 mya. Within the Colorado River Basin, divergence occurred 1.9-1.3 mya between the upper and lower

basin populations. Once these populations diverged, morphological adaptations were arose for each region based on the current gradient that existed in each habitat (Smith & Dowling, 2008).

R. osculus is the most frequently occurring freshwater fish in the state of Oregon. Pfrender, Hicks, & Lynch (2004) examined a 670 bp segment of cytochrome b in mtDNA which revealed deep divergence and genetic variation among basins. When researchers applied Kimura's molecular clock hypothesis to their data, they were able to attribute the divergence in populations to vicariant events that occurred during the late Miocene and early Pliocene. Two additional mtDNA lineages were discovered to co-occur in the Klamath River Basin which may reflect the isolation and then re-connectivity of the populations within the basin, and provides some insight into the population substructure that exists in some of the Northern Californian populations (Pfrender, Hicks, & Lynch, 2004).

Recently, Hoekzema & Sidlauskas (2014) identified cryptic species of Speckle Dace in the Oregon Great Basin region utilizing both phylogenetics and population genetics. They utilized mtDNA's ND2 region and nuclear DNA's S7 sequence data to further analyze the phylogenetic relationships of the various Speckle Dace populations within the Oregon Great Basin region. Eight microsatellites were used to examine the population genetics of each population. They were able to identify 3 distinct clades within the region and high levels of population structure occurring in an isolated spring (Hoekzema & Sidlauskas, 2014).

One of the largest river basins in the Southern California region is the Santa Ana River watershed, covering 6,900 km^2 in four different California counties including San Bernardino, Riverside, and Los Angeles (Santa Ana River Watershed Project Authority, 2004; Santa Ana Watershed Association, 2009). The watershed's topography is highly variable with high elevation peaks in the north and east and more arid, semi-desert conditions found to the west along the coastal plains (Santa Ana Watershed Association, 2009).

The Santa Ana Watershed can be divided into the Upper and Lower Watershed. The Upper Watershed is located between San Gorgonio Peak (San Bernardino Mountains) and the Prado Basin, whereas the Lower Basin consists of the regions below Prado Basin to the Pacific Ocean. Within each watershed are tributaries which are habitat for *R. osculus*. Santiago Creek is the only major tributary within the Lower Watershed that previously contained populations of *R. osculus* but those populations are extirpated (Santa Ana River Watershed Project Authority, 2004; Santa Ana Watershed Association, 2009). Indian Creek is also part of the Santa Ana Watershed and is located in the San Jacinto Mountains of the Peninsular Range.

R. osculus carringtonii is the local subspecies that was identified within the Santa Ana River Watershed by Cornelius (1969) based on morphological measurements. He concluded that *R. osculus carringtonii* was more closely related to the populations of *R. osculus* yarrow from Arizona's Virgin River and the Colorado Basin, rather than to populations of *R. osculus carringtonii* found in

Northern California. This was further supported by the research of Oakey, Douglas, & Douglas (2004) suggesting that that Colorado Basin population diverged approximately 1.9 mya and gave rise to the Los Angeles Basin populations.

The Santa Ana River Watershed has gone through many geological and locational changes throughout history. During the end of the last glacial period, the Wisconsinian Glaciation event, many of the rivers within Southern California flooded due to the melting of glacial ice (Colburn, 2006). It was during this time that the Santa Ana River began eroding away the granite formations of the surrounding mountains, including the Peninsular Ranges. As a result of this flooding, many of the tributaries, which would normally be isolated, came into connection with one another. It was during these kinds of events that researchers (Pfrender, Hicks, & Lynch, 2004; Smith & Dowling, 2008) believe connectivity between populations occurred. Once isolation retuned, the populations eventually became genetically structured. *R. osculus* would prove to be a valuable tool for examining population structure among and within tributaries.

The Santa Ana Speckled Dace, one of the rarest native freshwater fish in Southern California, occupies only remnants of its native range (Figure 2) as a result of anthropogenic effects such as urbanization developments and county controlled flood measures. The population within the Los Angeles River system were extirpated in the early 1990's (Santa Ana Watershed Association, 2009;

Santa Ana River Watershed Project Authority, 2004). The populations within the Santa Ana and San Gabriel watersheds are in imminent danger of extirpation. These populations have a very limited range and can be greatly affected by fires and floods (Moyle, Yoshiyama, Williams, & Wikramanayake, 1995).

Government agencies have listed the Santa Ana Speckle Dace as a species of concern in both 1995 (California Department of Fish and Wildlife) and 1998 (United States Forest Service). The Santa Ana Speckled Dace never received federal protection under the Endangered Species Act due to the lack of peer reviewed descriptions of their taxonomic status (Moyle & Marchetti, 2006).

The fires and floods that occurred within the Santa Ana Watershed during 2003-2004 impacted the geographic distribution of *R. osculus*. Prior to the fires and floods, the five northernmost drainages that drain into the Santa Ana River had recorded population of *R. osculus*. These drainages included Lytle Creek, Cajon Creek, Twin/Strawberry Creek, City Creek and Plunge Creek. After the fire and floods occurred, the populations in Twin/Strawberry Creek and City Creek were extirpated. If agencies such as California Department of Fish and Game or the United States Forest Service decide to reintroduce populations back into areas in which *R*. *osculus* was extirpated or to introduce them into new suitable habitats, an analysis of the genetic variation that occurs between drainages as well as within drainage needs to be performed. This will allow for the stock population to contain a suitable amount of genetic variation to sustain a viable population.

The primary focus of this study is to identify and characterize microsatellite loci that can be used to characterize the genetic variation present within and among the populations of *R. osculus* inhabiting the Santa Ana and San Gabriel River watersheds of Southern California. Utilizing the microsatellite data, the degree of gene flow will be examined to determine if the current lack of connectivity between the drainages has caused each tributary to show patterns of genetic differentiation and population structure. In addition, the Santa Ana Speckled Dace populations will be compared to populations located in the Owens Valley region and those found on the Central Coast of California to provide further insight into the patterns of genetic differentiation of the Santa Ana Speckled Dace. I hypothesize that there is a high level of genetic structure that occurs in the speckled dace. Current mtDNA in Tony Metcalf's lab has shown a 6.33-6.78% genetic difference between the Central Coast and Owens Valley regions to that of Southern California. This suggests independent evolutionary trajectories. Microsatellite loci will be used to further examine this regional structure. I hypothesize that a high level of genetic structure will occur among the three regions due to the lack of gene flow and isolation by distance between each of the three regions.

In this study, specimens of *R. osculus* were obtained from various sample locations throughout the Santa Ana and San Gabriel River Watershed to examine population substructure among and within tributaries in the Southern California region. Previous work has been done our lab to examine the extent of

genetic variation that occurs among mtDNA's *cytochrome b* and *d-loop* region, but this provides only a maternal perspective of the evolutionary lineage for *R. osculus* in the Southern California region. Genetic differentiation results from the isolation of a population and the forces that act upon that population resulting in unique genetic structures. Using microsatellite loci genetic diversity and population structure were examined for the Santa Ana Speckled Dace populations providing a complete genetic description of *R. osculus* at the population and individual level, for this region.

The Southern California region has gone through extensive geological and climatal changes throughout its history. This has altered the landscape and topography of the region, during pluvial and arid intervals (Colburn, 2006). Each of these events could alter the evolutionary history of a species by influencing the levels of dispersal as well as gene flow that can exist between populations. Freshwater tributaries within the Santa Ana River Watershed have also undergone such events that which can lead to the divergence of aquatic populations. *R. osculus* was once able to freely inhabit the free flowing and perennial rivers within the Santa Ana River Watershed, but due to changes in climate and geology and more recent anthropogenic causes in the last century, their habitat became discontinuous and fragmented. Currently, due to the discontinuous and fragmented habitats of *R. osculus*, the degree of gene flow between population is very limited, if it exists at all. (Cornelius, 1969; Oakey, Douglas, & Douglas, 2004) I therefore expect to see evidence of historical

admixture when connectivity between the tributaries was possible, but as a result of current anthropogenic effects leading to fragmented habitats, the populations should exhibit a moderate degree of structure as a result of their isolation.

Through my analysis I hope to gain a better understanding of the molecular evolution, phylogeography, population genetics and conservation of the local populations of the Santa Ana Speckled Dace.

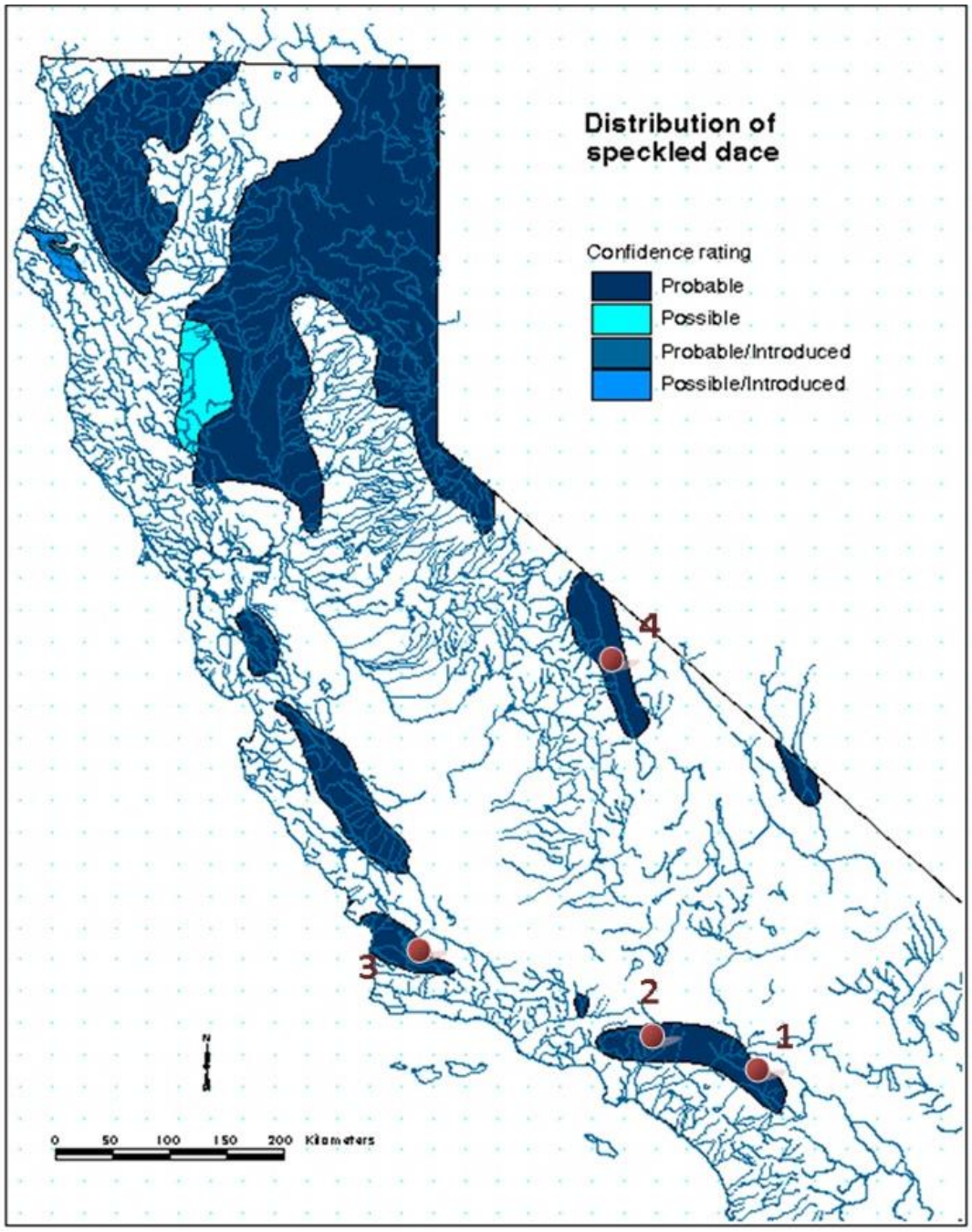


Figure 1. Distribution of the Speckled Dace within California. Areas bordered in dark are known to contain Rhinichthys osculus within river drainages of that region. Watershed sampling locations are denoted in red numbered circles. 1=Santa Ana River, 2=San Gabriel River; 3=Santa Maria River and San Luis Obispo River; 4=Owens River. (Map adapted from Nico, L., & Fuller, P. (2015). *Rhinichthys osculus*. Retrieved from USGS Nonindigenous Aquatic Species Database: http://nas.er.usgs.gov/)

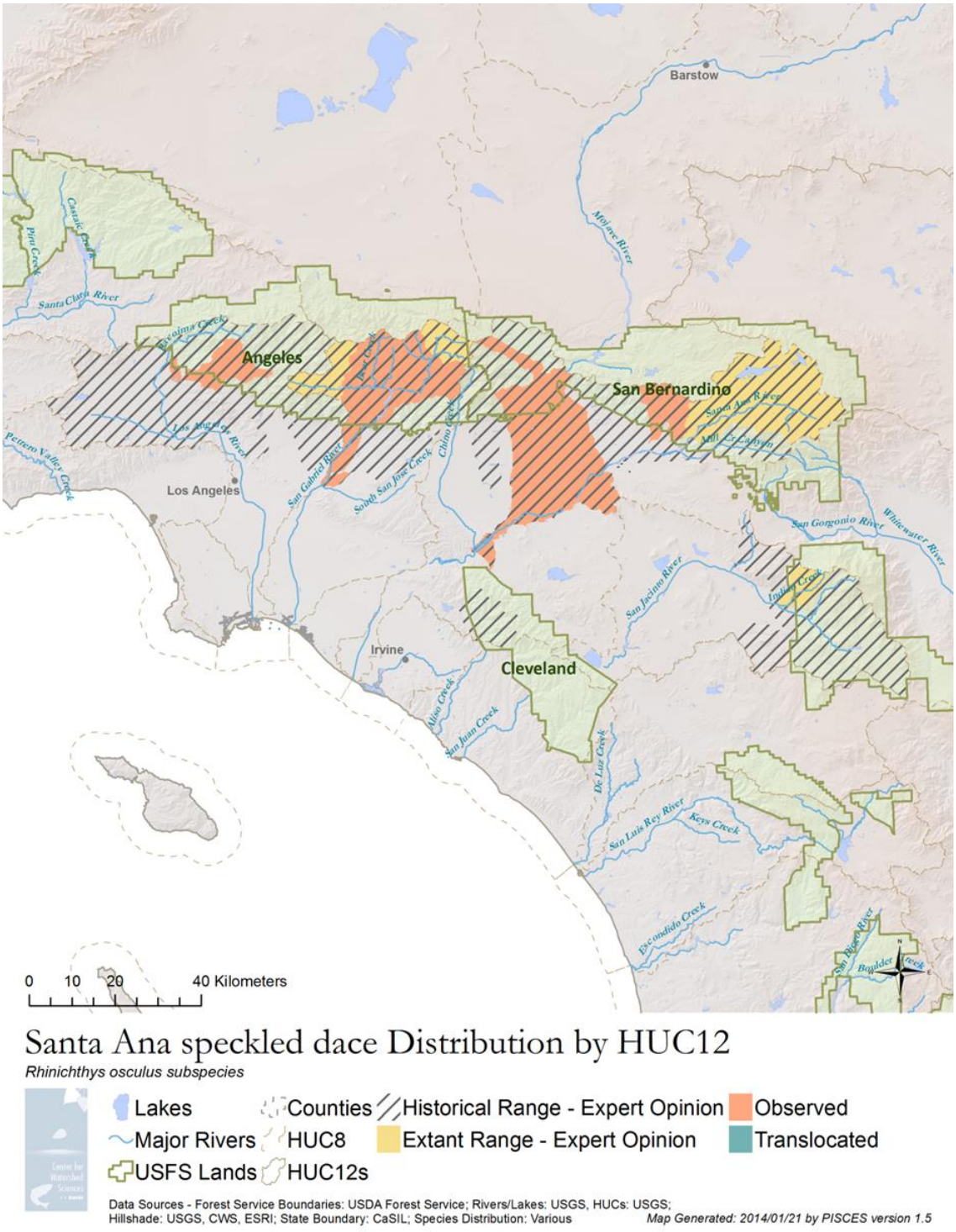


Figure 2. Current and Historical Distribution of the Santa Ana Speckled Dace (R. osculus) Map taken from (Santos, N. R., Katz, J. V., Moyle, P. B., & Viers, J. H. (2014). *RHINICHTHYS OSCULUS SUBSPECIES.* Retrieved from UC DAVIS PISCES: http://pisces.ucdavis.edu/content/rhinichthys-osculus-subspecies-2).

CHAPTER TWO

MATERIALS AND METHODS

Specimen Collection

In order to evaluate patterns of genetic variation that exist between the various tributaries within the population inhabiting Southern California, the Central Coast of California and the Owen's River Valley, samples of *Rhinichthys osculus* were collected from various sampling sites as illustrated by Figure 1. Samples were acquired in conjunction with the United States Forest Service and California Department of Fish and Wildlife. Additional samples were collected using proper electroshocking techniques in accordance with permits issued to the Metcalf Lab by the United States Forest Service at various sampling locations. For each specimen that was collected by the Metcalf Lab, GPS coordinates were taken from the sampling location on the designated tributary. A minimum of seven *R. osculus* individuals were collected from each tributary. *R. osculus* sampling sites within the Santa Ana Watershed (Southern California population) included Plunge Creek, City Creek, Twin/Strawberry Creek, Cajon Creek, Lytle Creek, Mill Creek, and Indian Creek (Figure 2). The San Gabriel River Watershed (Southern California population) included sampling sites within Cattle Canyon Creek, Fish Creek, North Fork, West Fork, and East Fork of the San Gabriel River (Figure 2). In addition, two samples were obtained from the Hain's River located within the Los Angeles River Watershed (Southern California population) (Figure 2). *R. osculus* specimens representing the Central Coast

population were collected in two different watersheds, the San Luis Obispo Watershed and the Santa Maria Watershed. The San Luis Obispo watershed included specimens collected from San Luis Obispo Creek, Stenner Creek, and Brizziolari Creek (Figure 1). Tributaries within the Santa Maria Watershed, in which R. osculus specimens were collected, was composed of the Cuyama River, Manzana Creek, and Davy Brown Creek (Figure 1). *R. osculus* samples representing the Owen's River Watershed were obtained from Marvin's Marsh and Pine Creek (Figure 1). A list of all samples (n=146) and locations are represented in Table 1.

Molecular Methods

DNA was isolated using phenol-chloroform extraction methods utilizing phase lock gels (PLG) as set forth in Eppendorf's Phase Lock Gel Manual (Mouse Tail Genomic DNA Isolation Protocol) or Qiagen DNeasy kits. For phenol-chloroform extractions using Phase Lock Gel tubes, a 10-50 mg tissue sample was extracted from each specimen and then washed three times to remove ethanol. To further extract ethanol from the tissue, samples were placed under vacuum conditions and left for 90 minutes. Each sample was then placed in a 1.5 milliliter centrifuge tube and digested with a lysis Buffer (50 mM Tris-HCl, pH 8.0, 100 mM EDTA, 100 mM NaCl, 1% SDS) and proteinase-K in 55°C water bath for at least 12 hours. Genomic DNA was then extracted from the solution using 2 phenol-chloroform-isoamyl (25:24:1) and 1 chloroform-isoamyl (24:1) extractions in Phase Lock Gel (Heavy) tubes. Further treatments of the solution

with sodium acetate and 95% ethanol were utilized. The tubes were then centrifuged to create a genomic DNA pellet in which the remaining solution was then decanted. Samples were air dried for at least 45 minutes and then each DNA extraction was suspended in 150µL of TE Buffer (1mM Tris-HCL, 1mM EDTA, pH 8.0). Genomic DNA extractions were then visualized using agarose gel electrophoresis. Using the Qiagen DNeasy kits, a 10-25 mg piece of tissue was removed from each whole sample specimen. The tissue was then rinsed three times and placed under vacuum conditions for at least 90 minutes in order to fully extract the ethanol from the tissue sample. The sample was then transferred to a sterile 1.5 milliliter microcentrifuge tube where Qiagen's Buffer ATL and proteinase-K were added to the tube in order to digest the tissue sample. The tube was then placed in a 55°C water bath for at least 12 hours. Qiagen's Buffer AL and 95% ethanol were then added to the tube and centrifuged. The solution was then transferred to a Qiagen spin column and collection tube, which was further treated with additional rounds of centrifugation with Qiagen Buffer AW1 and Buffer AW2. Spin columns were transferred to new collection tubes after each round of centrifugation. The spin column was then transferred to a 1.5 milliliter microcentrifuge tube where Qiagen's Buffer AE was added and then the tube was centrifuged; resulting in the extraction of the specimen's genomic DNA. DNA concentrations were analyzed using spectrophotometry (A260/A280 and A234/A260) to ensure the purity and yield of the samples.

Microsatellite loci identification and isolation as well as primer development was performed in conjunction with the Savannah River Ecology Lab. 24 *R. osculus* specimens were used to create an Illumina pair-end shotgun library in which five million successful reads were used to identify 4635 nuclear base pair repeats. Primers were created for these microsatellite loci and they were screened to identify sequences that would occur only once; identifying 48 loci. Eight *R*. *osculus* samples were used to examine the polymorphic characteristics of the loci in which 23 microsatellite loci appeared as prospective candidates. To further access the reliability and polymorphism of the microsatellite loci, the 23 microsatellite loci were tested across all 24 specimens encompassing various watersheds within California (Nunziata, Lance, Jones, Nerkowski, & Metcalf, 2013). Table 2 summarizes the 23 microsatellite loci that were identified for *R. osculus*.

Once the microsatellite loci had been characterized and identified, each microsatellite locus underwent preliminary screening to determine which loci were to be used in this study. Following a protocol for visualizing microsatellites using unlabeled primer for the long nose dace on 2% metaphor gels (Girard & Angers, 2006), in which a subset of 40 of the 135 *R*. *osculus* samples were used for amplification of each of the 23 loci. Each 25µL PCR reaction consisted of 1.5µL of 1.5mM of MgCl2, 1µL of each 10µM dNTP solution, 0.5µL of Taq polymerase, 2.5µL of 10x Taq polymerase buffer, 2µL of each primer (5mM), 14.5µL dH2O, and 75ng of genomic DNA. The PCR program, utilizing BioRad

C1000 Thermal Cycler, consisted of an initial denaturing temperature of 92°C for 30s, then 45 cycles of the following profile: 92°C for 30 seconds for denaturation, 15 seconds at the annealing temperature 65°C, and 5 seconds at 68°C. The final phase of the cycle is a 2 minute extension at 68°C. PCR amplicons were then visualized utilizing a 2% METAPHOR© gel and analyzed as to the successfulness of the loci on the 40 samples with the current DNA extractions.

Each of the 146 samples of *R. osculus* underwent PCR amplification for seven of the 23 polymorphic microsatellite loci; for this study *Rhos 5, 8, 9, 14, 23, 25* and *29* were utilized based on reliability and repeatability. PCR conditions were optimized and carried out in 10µL volumes: 1.0 µL template DNA (approximately 20-40 ng/µL), Rhos *23*, *25*, and 29 used 0.6µL of each primer (1mM) (*Rhos8* and *14* used 0.5µL, *Rhos9* used 0.4µL, and *Rhos5* used 0.4µL), 0.2µL Taq polymerase, 1.0µL dNTP's (2mM), 5.0µL buffer (10x buffer with 15mM MgCl2), and sterile water (molecular grade) brought up each reaction to 10µL. The PCR profile for all reactions, across all seven loci was: 95°C for 5 minutes (initial denaturation), 36 cycles of denaturation at 95°C for 20 seconds, annealing at 65°C for 20 seconds, and extension at 72°C for 30 seconds; and a final extension of 72°C for 3 minutes. 3µL of LI-COR, Inc. Blue Stop Solution was added to each sample and then denatured using the PCR machine at 95°C for 3 minutes. Samples were then immediately placed on ice and then run on a 6.5% polyacrylamide gel using a LI-COR 4300 automated sequencer. Each sample was visualized on three different gels to ensure the accuracy of the microsatellite

scoring. Microsatellite images were visualized, analyzed and scored using SAGA GT software (LI-COR, INC.).

Population Genetics

The goal of this study was to characterize the genetic variation and population structure that occurs among and within the populations of speckled dace within the Southern California region and compare it to those speckled dace populations residing in the Owens River and Central Coast ranges. Within the Southern California populations, genetic variation and population structure were analyzed using four different models: each tributary, between watersheds (Santa Ana and San Gabriel), between mountain ranges, and isolation by distance. The analysis will provide insight into the current and historical gene flow occurring in the Southern California region. In addition, analyzing the microsatellite data for each region within California may provide further insight into the regional population level structure and provide a more complete genetic characterization of the Santa Ana Speckled Dace.

Genetic diversity indices were calculated for each region, watershed, mountain range, and tributary using GENALEX version 6.501 (Peakall & Smouse, 2006;Peakall & Smouse, 2012). This included number of alleles per locus (N_A), observed (H_O) and expected heterozygosity (H_E), allelic richness (AR), percent polymorphic loci, and the number of private alleles.

Genetic differentiation can be examined through an F-statistic created by Wright (1931) known as the F_{ST}. F_{ST} values can range from 0 to 1, where

increasing values represent higher degrees of genetic differentiation. The F_{ST} value compares the level of genetic variation that occurs within the subpopulations and compares it to the level in the overall population (Hartl & Clark, 2006; Wright, 1931). Analogs of F-statistics such as R_{ST} values (Balloux & Goudet, 2002) have been developed to analyze population structure utilizing microsatellite loci; this model accounts for step-wise mutations which are characteristic of such loci. R_{ST} values were examined in GENALEX version 6.501 (Peakall & Smouse, 2006;Peakall & Smouse, 2012). Pairwise F_{ST} values were calculated in GENALEX following Weir & Cockerham (1984).

AMOVA, analysis of molecular variance, is a statistical model that examines patterns of molecular variation within and among population using F-statistics or its analogs (Excoffier, Smouse, & Quattro, 1992; Excoffier & Slatkin, 1995; Excoffier, Laval, & Schneider, 2005). GENALEX version 6.501 was used to run AMOVAs. Four different AMOVA's were implemented; the first three examined the patterns of genetic variation that occurred within the Southern California region (by tributary, watershed, and mountain range) and the fourth examined the patterns of genetic variation and differentiation between the regions within California. Individuals were assigned to populations in the first AMOVA by the tributary in which they were collected; where each tributary within the Santa Ana and San Gabriel River Watersheds were considered independent populations. Individuals, for the second AMOVA, were assigned to populations based on the watershed in which the tributary was located (Santa Ana vs. San

Gabriel River Watershed in which Hain Creek was considered part of the San Gabriel River due to the small sample size). Individuals for the third AMOVA were assigned to populations based on the mountain region in which the headwaters of the tributary was located (San Jacinto, San Bernardino or San Gabriel Mountains). Finally, the fourth AMOVA assigned the individuals to populations based on the region of collection (Southern California, Central Coast, and Owen's River populations). Significance testing of F-statistic values and variance components were based on 999 permutations and evaluated at the 0.05 probability level.

A Mantel test was performed in GENALEX version 6.501 to examine the correlation between the pairwise F_{ST} and R_{ST} values from the microsatellite data, and geographic distance (km) between populations or regions. Pairwise geographic distances (km) were determined using centroid calculations. Significant results from the Mantel test indicates isolation by distance among the populations sampled.

Patterns of genetic variability and inferred populations (K) were identified using STRUCTURE (version 2.3.4) (Falush, Stephens, & Pritchard, 2003; Falush, Stephens, & Pritchard, 2007, Hubisz, Falush, Stephens, & Pritchard, 2009; Pritchard, Stephens, & Donnelly, 2000) K will indicate the number of populations that are best represented based on the genetic variation and similarities in the haplotypes of the microsatellite data. In addition, STRUCTURE will be used to determine the proportion of each individual's genome that arose

from the inferred populations. A parameter was set in which the admixture model was used, with correlated allele frequencies, and LOCOPRIOR (prior location of capture). Under this parameter, 15 independent runs of K = 1-10 (where K represents the number of putative clusters) with an initial burn-in period of 100,000 iterations and a run length of 1,000,000 Markov chain Monte Carlo iterations was performed. STRUCTURE HARVESTER 6.92 (Earl & vonHoldt, 2012) was used to assess the most probable number of clusters represented by the data using the second order rate of change, ΔK, and LnP(D) under the Evanno, Regnaut, & Goudet (2005) method (Appendix C). Utilizing the 15 iterations of the most probable K, data was inputted into CLUMPP 1.1.2 (Jakobsson & Rosenberg, 2007) using a full-search algorithm. Graphical representation of the STRUCTURE results was created using DISTRUCT (Rosenberg, 2004)

To further examine clusters of genetically similar individuals, a discriminate analysis of principal components (DAPC) was performed in R (version 3.2.0) package *adegenet* (Jombart & Ahmed, 2011). A DAPC does not assume a particular underlying model of gene flow as is used in a principal component analysis (PCoA)

Table 1. *Rhinichthys osculus* Sampling Locations. (n) represents sample size at each location; Mountain Range codes are as follows: San Bernardino (SB), San Gabriel (SG), San Jacinto (SJ), Coastal Ranges (CR), and Eastern Sierra Nevada (SN). Regions are coded as Southern California (SOCAL), Central California (CENTRAL), and Owen's Basin (OWENS).

POPULATION	n	WATERSHED	MOUNTAIN RANGE	REGION
CAJON CREEK	21	SANTA ANA	SB	SOCAL
CITY CREEK	14	SANTA ANA	SG	SOCAL
LYTLE CREEK	28	SANTA ANA	SG	SOCAL
INDIAN CREEK	7	SANTA ANA	SJ	SOCAL
MILL CREEK	5	SANTA ANA	SB	SOCAL
PLUNGE CREEK	21	SANTA ANA	SB	SOCAL
TWIN CREEK	12	SANTA ANA	SB	SOCAL
FISH CREEK	5	SAN GABRIEL	SG	SOCAL
CATTLE CANYON	3	SAN GABRIEL	SG	SOCAL
NORTH FORK SAN GABRIEL RIVER	2	SAN GABRIEL	SG	SOCAL
EAST FORK SAN GABRIEL RIVER	2	SAN GABRIEL	SG	SOCAL
WEST FORK SAN GABRIEL RIVER	1	SAN GABRIEL	SG	SOCAL
HAIN CREEK	2	LOS ANGELES	SG	SOCAL
PINE CREEK	3	OWENS	SN	OWENS
MARVIN'S MARSH	3	OWENS	SN	OWENS
BRIZZOLARI CREEK	3	SAN LUIS OBISPO	CR	CENTRAL
STENNER CREEK	3	SAN LUIS OBISPO	CR	CENTRAL
SAN LUIS OBISPO CREEK	3	SAN LUIS OBISPO	CR	CENTRAL

POPULATION	n	WATERSHED	MOUNTAIN RANGE	REGION
DAVY BROWN CREEK	2	SANTA MARIA	CR	CENTRAL
CUYAMA RIVER	2	SANTA MARIA	CR	CENTRAL
MANZANA CREEK	4	SANTA MARIA	CR	CENTRAL

Table 2: Microsatellite Loci and Primer Information Developed in Conjunction with the Savannah River Ecology Lab. The size (bp) indicates the range of observed alleles in base pairs and includes the length of the CAG tag; TA refers to the annealing temperature where TD65 indicates a touchdown protocol and number of individuals genotyped is N. * indicates CAG tag (5'- CAGTCGGGCGTCATCA-3') label. Loci bolded were used in genotyping the 146 speckled dace samples (Modified from Nunziata, S. O., Lance, S. L., Jones, K. L., Nerkowski, S. A., & Metcalf, A. E. (2013). Developement and characterization of 23 microsatellite markers for Rhinichthys osculus using paired-end Illumina shotgun sequencing. *Conservation Biology Notes*.)

Locus	**Primer Sequence 5' --> 3'**	**Repeat motif**	**Size (bp)**	**TA**
Rhos1	F: *TTAAATTGTGCCAAATGATGC R: AAGACACATTTGTTTGGAAGGC	AATG	228-236	65
Rhos3	F:*TCAGCTAACCAAATATTGCATGG R:ACAAACGGGAAGGAGCAGG	ATCT	268-436	65
Rhos5	F:*TGGCATTGAGCGAGGTCC R:ACGATATTTAGCTGTCATCATCCG	ATCT	237-313	65
Rhos8	F: *TCGCAAAGATTCACAAACGG R: ATCAGCTCACAATGATCCGC	AAAG	173-281	65
Rhos9	F: CGAGAATGACTCAACATTAACC R: ATGTTGGCACGTGAAAGCC	ATCT	145-185	65
Rhos10	F:*TTGGACAGCTGTATGAATTGGG R:ATTGCAGGACCACCAACACC	ATCT	305-377	65

Locus	Primer Sequence 5’ --> 3’	Repeat motif	Size (bp)	TA
Rhos14	F: *GGTGCAGCTTTGAGAGGG R: CCTATATTAAACTCTATGAGCCATAAATCC	ATCT	262-322	65
Rhos16	F:* TGTTACTAATAATCATGTCCTGAAGAGG R: CGCTACTCTGGGTTTGAATGC	AATG	168-312	65
Rhos18	F: *AACTATAACCAGGTGTTACAGTGGG R: CGTAGTACAATGTTACACAATAATAGGC	ATCT	252-360	TD65
Rhos20	F:*GAGGACTGTTTCTATCCCGGC R:CTGGAAATCACAAACCAGGG	ATCT	246-366	65
Rhos21	F: *CCTACAATGTTTGTTTGTTTACACG R: CTTTGAGGAGATTAACTCTTTCCC	AAAG	162-246	65
Rhos22	F: *CAATGTTTCCAATTCTATAACAAAGG R: GAAGACATCAGACATCCAATTTCC	ATCT	365-413	65
Rhos23	F:*TGTGTAAACCAGTAGACTTTCTAATATACC R:GACAATGAAACAATGATTACTTACAGC	ATCT	218-358	65
Rhos25	F: *TGTCAGTATCAGTCTTTCCTTCGC R: GAGCCATATGCATGGAGAGC	AAAG	126-214	65
Rhos26	F: *TGGAATATATCCAGTACCAACTTCC R: GATGGACATAGAACAATGGATGG	ATGG	103-163	65
Rhos27	F:*AAACTGTAACCTAAATGGGTATTAAAGAGG R: GCATGAAACGATGTGAAATAGCC	ATCT	274-330	65

Locus	Primer Sequence 5’ --> 3’	Repeat motif	Size (bp)	TA
Rhos29	F: *TCTCAAGCTCTTATTCTGATGATCC R: GCTTACTGTGCTGTCTGTGGC	ATCT	204-264	65
Rhos31	F: *TCATACTGCCGTCTAGTGGTGG R: GTCATCGGGTCAGCAGAGG	ATCT	173-305	65
Rhos33	F:*AGGTGATGCCACACATGACG R:TCAAAGAATCTGAGCGTCGG	ATCT	324-372	65
Rhos35	F:*TGTTCAACAGGCCTCAAACC R:TCATTACCTTATTAAAGGGACAGTGC	ATCT	302-354	65
Rhos36	F: *TGTTCAACAGGCCTCAAACC R: TCCCTTTATACTTTCAGCTGCTCC	AAAG	144-204	TD65
Rhos42	F: *AAATGAGCAAGTGAGCCAGC R: TGATCATTAGGAAGGATACACTGC	AAAC	140-191	TD65
Rhos43	F: *AGTGGAACATCAGTCACTGCG R: TGATGATATGTGCATCAAGCG	AAC	256-274	TD65

CHAPTER THREE

RESULTS

Genetic Diversity

Population genetic analysis was used to assess the levels of genetic differentiation within and among the Speckled Dace populations of California. Genetic indices including the number of alleles examined (N), alleles for each locus (N_A), observed (H_O) and expected (H_E) heterozygosity and the Fixation Index (F_{IS} = (He - Ho) / He = 1 - (Ho / He)) for each of the seven microsatellite loci examined are presented in Table 1. *Rhos5* had a range of 220-308 bp in which 21 alleles were identified among the 146 *R. osculus* samples. *Rhos8* exhibited a range between 160-298 bp, 26 different alleles were identified; *Rhos9* exhibited a range of 117-225 bp, 22 alleles were identified; *Rhos14's* exhibited a range was 230-314 bp, 18 alleles were identified; *Rhos23* exhibited a range of 196-340 bp, 24 alleles identified; *Rhos25* exhibited a range of 108-220 bp, 26 alleles identified; and *Rhos29* exhibited a range of 175-291 bp, 20 alleles were identified.

Mean heterozygosities across the seven loci ranged from 0.333 (Brizzolari) to 1.000 (East Fork and North Fork San Gabriel River). Significant deviations from Hardy-Weinberg expectations occurred in five populations, with two or three of the microsatellite loci analyzed, within all the sampled regions.

Loci deviating from Hardy-Weinberg proportions occurred within the Santa Ana River Watershed only (Table 3).

Table 5 summarizes each tributary's genetic indices including allelic richness (AR), observed (H_O) and expected (H_E) heterozygosity, and the number of private alleles. Private alleles were identified in all of the tributaries within the Southern California populations except for Mill (n=5), Cattle Canyon (n=3) and West Fork of the San Gabriel River (n=1). This may simply be due to the sample size obtained from these populations. The greatest number of private alleles was observed in Plunge Creek

Table 3. Speckled Dace Populations Deviating from Expected Hardy-Weinberg Proportions

POPULATION	LOCUS	p-VALUE
INDIAN CREEK	Rhos25	0.021
PLUNGE CREEK	Rhos14	0.009
PLUNGE CREEK	Rhos23	<0.001
TWIN CREEK	Rhos25	0.003
TWIN CREEK	Rhos29	0.019
CITY CREEK	Rhos5	0.008
CITY CREEK	Rhos23	<0.001
CAJON CREEK	Rhos9	0.03
CAJON CREEK	Rhos25	0.012
LYTLE CREEK	Rhos8	0.007
LYTLE CREEK	Rhos9	0.007
LYTLE CREEK	Rhos25	<0.001

Statistics are based on Chi-Square Tests for Hardy-Weinberg Equilibrium. Only loci listed are display significant deviations from expected proportions. All tributaries deviating from Hardy-Weinberg proportions are located in the Southern California region.

Table 4. Genetic Diversity for the Southern California, Owens Valley, and Central Coast Populations of Speckled Dace.

POP		Rhos 5	Rhos 9	Rhos 25	Rhos 23	Rhos 14	Rhos 8	Rhos 29
SOCAL	N	123	123	123	123	123	123	123
	N_a	19	13	24	17	16	20	18
	N_e	11.540	5.227	11.295	8.832	7.610	10.212	7.378
	H_o	0.805	0.618	0.626	0.520	0.650	0.724	0.724
	H_e	0.93	0.809	0.911	0.887	0.869	0.902	0.864
	F	0.119	0.236	0.313	0.413	0.251	0.198	0.163
OWENS VALLEY	N	6	6	6	6	6	6	6
	N_a	8	8	2	9	7	9	8
	N_e	6.000	5.538	1.385	8.000	6.000	7.200	7.200
	H_o	0.833	0.667	0.333	1.000	0.500	1.000	0.333
	H_e	0.833	0.819	0.278	0.875	0.833	0.861	0.861
OWENS VALLEY	F	0.000	0.186	-0.200	-0.143	0.400	-0.161	0.613
CENTRAL COAST	N	17	17	17	17	17	17	17
	N_a	13	10	3	15	4	1	12
	N_e	8.892	5.838	1.966	7.811	1.357	1.000	6.964
	H_o	0.765	0.529	0.353	0.647	0.118	0.000	0.412
	H_e	0.888	0.829	0.491	0.872	0.263	0.000	0.856
	F	0.138	0.361	0.282	0.258	0.553	#N/A	0.519

Genetic diversity includes number of samples for each region (N), number of alleles present for each loci (N_a), number of effective alleles (N_e), observed (H_O) and expected (H_e) heterozygosities for each region, and fixation index (F).

Table 5. Santa Ana Speckled Dace Genetic Diversity Summaries For Each of the Tributaries in the Southern California Region. Genetic diversity summaries for each of the 12 tributaries sampled in the Southern California region where the number of samples (n), average allelic richness across all seven loci (AR), observed (Ho) and expected (H_E) heterozygosities, and private alleles (PA) are identified.

POPULATION	n	AR	H_0	H_E	PA
CAJON CREEK	21	9.714	0.782	0.827	2
CITY CREEK	14	8.429	0.673	0.780	2
INDIAN CREEK	7	4.143	0.490	0.536	1
LYTLE CREEK	28	5.429	0.571	0.626	2
MILL CREEK	5	4.286	0.743	0.634	0
PLUNGE CREEK	21	8.714	0.673	0.825	6
TWIN CREEK	12	6.857	0.595	0.711	1
CATTLE CANYON	3	4.857	0.905	0.754	0
FISH CREEK	5	4.714	0.600	0.651	1
NF SAN GABRIEL	2	3.143	1.000	0.643	1
EF SAN GABRIEL	2	3.143	1.000	0.643	2
WF SAN GABRIEL	1	1.714	0.714	0.357	0
HAIN RIVER	2	2.286	0.643	0.500	1

Population Genetic Structure

Geographic structuring among the three regions sampled was evaluated in STRUCTURE HARVESTER utilizing the STRUCTURE runs for K=1-10. The most probable K was identified as K=3 when examining the cluster arrangement between the three regions sampled in California. All STRUCTURE runs of K=3 produced similar membership coefficients (q scores). Geographic subdivision revealed three groups, Central Coast, Southern California and Owens Valley (Figure 8a). The results of the DAPC corroborated the results of STRUCTURE where K=3 (Figure 7) was inferred in which the Southern California, Central Coast and Owens River populations clustered independently of one another, supporting the genetic differentiation observed between the three regions. Each sample was properly assigned to the region it was collected based on the microsatellite data analysis.

The greatest amount of genetic differentiation was observed when comparing the three regions within California (R_{ST}=0.600, $p<0.05$). 60% of the variation observed in the microsatellite data occurred among the three regions whereas differences within populations accounted for only 40% of the variance (Table 6 and Table 7b). Pairwise F_{ST} values from the AMOVA analysis are summarized in Tables 7a and 8a. A Mantel test found a significant correlation (F_{ST}: R^2=.1808, p=0.010) between population pairwise F_{ST} values and geographic distance (km), suggesting isolation by distance between the three regions of California (Figure 3). In addition, when accounting for the stepwise

mutation model exhibited by microsatellites, a Mantel test found a significant moderate correlation between pairwise R_{ST} values and geographic distance (R^2=0.361, p=0.010; Figure 4).

Table 6. R_{ST} and N_M values for Analysis of Molecular Variance Comparisons

COMPARISON	R_{ST} VALUES	p-VALUES	N_M
Among all tributaries within Southern California	0.160	0.001	1.373
Between Santa Ana and San Gabriel River Watersheds	0.151	0.001	1.447
Between San Jacinto, San Bernardino and San Gabriel Mountain Ranges	0.077	0.001	3.134
Between Southern California, Central Coast and Owens Valley	0.600	0.001	0.168

R_{ST} values and their significance levels for an overall AMOVA and then AMOVA's analyzing different models that would explain population structure within and among the Santa Ana Speckled Dace populations.

Further geographic structuring was evaluated for the Southern California populations. STRUCTURE runs of K=1-10 were analyzed in STRUCTURE HARVESTER where the most probable K was identified as K=2; in which the Lytle Creek population was appeared to represent a single inferred population genotype and the remaining populations in Southern California represented an admixture of both of the inferred populations (Figure 8b). The greatest amount of

admixture with the Lytle population was observed in the most proximal tributary, Cajon Creek, and decreased with geographic distance. The Mantel test revealed a moderate correlation between geographic distance and genetic structure, suggesting isolation by distance is a contributing factor to differentiation among the Southern California populations (F_{ST}: R^2=0.1808, p=0.010 (Figure 5); R_{ST}: R^2=0.361, p=0.010(Figure 6)).

Results from the AMOVA analysis indicated significant levels of genetic differentiation occurred in all scenarios examined. The lowest R_{ST} values occurred when analyzing tributaries within Southern California were assigned to populations based on the mountain range in which they reside (R_{ST}=0.070; p<0.05). The majority of the variation, 92.3%, is occurred within the populations and not among them. When analyzing each tributary as its own population, moderate levels of population structure was observed (R_{ST}=0.160; p<0.05). When comparing the tributaries based on watersheds provided similar R_{ST} values (R_{ST}=0.151, p<0.05). In each analysis, 15-16% of the total variation occurred among tributaries or watersheds (Table 6). The remaining 84-85% of the variation could be accounted for within the tributary or watershed, suggesting moderate levels of population genetic differentiation.

Gene flow between speckled dace populations can be examined through the migration rate (N_M). Table 6 illustrates the number of migrants per generation for each of the comparisons utilized in the AMOVA scenarios. Under assumptions of equilibrium, the greatest number of migrants is observed among

mountain range model (N_M=3.134) and the weakest migration rate occurs between the three regions sampled within California (N_M=0.168), suggesting that these regions had little to no gene flow occurring historically.

Table 7. (a) AMOVA Pairwise R_{ST} Values Among the Southern California, Owens Valley and Central Coast Regions and (b) AMOVA Pie Graph Representing the Percentage of Molecular Variation Among and Within Sampled Regions.

(a)

	SOCAL	OWENS	CENTRAL
SOCAL	0.000	**0.001**	**0.001**
OWENS	0.406	0.000	**0.001**
CENTRAL	0.659	0.412	0.000

Pairwise RST values are displayed below the diagonal and bolded p-values represent significant values based on 999 permutations above the diagonal.

(b)

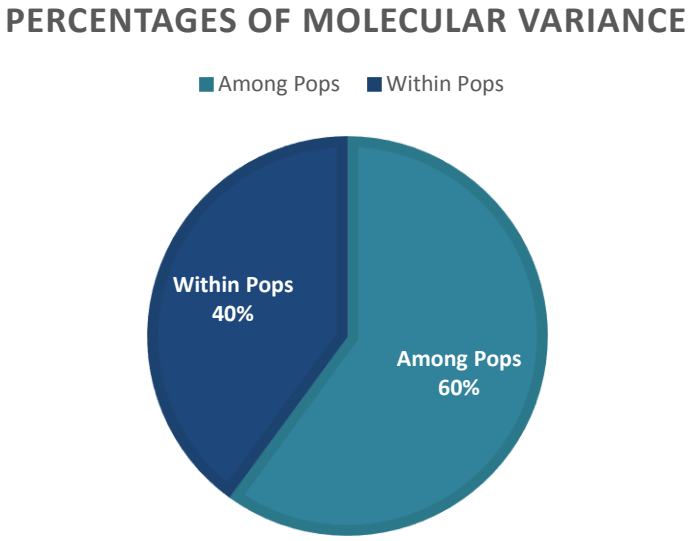


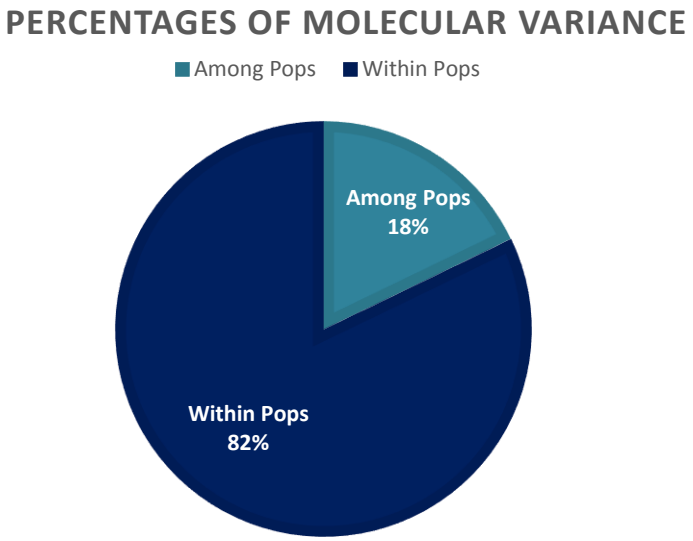
Table 8 (a) AMOVA Pairwise F_{ST} Values Among the Southern California, Owens Valley and Central Coast Regions and (b) AMOVA Pie Graph Representing the Variation Among and Within Sampled Regions.

(a)

	SOCAL	OWENS	CENTRAL
SOCAL	0.000	**0.001**	**0.001**
OWENS	0.090	0.000	**0.001**
CENTRAL	0.197	0.260	0.000

Pairwise FST values are displayed below the diagonal and bolded p-values represent significant values based on 999 permutations above the diagonal.

(b)



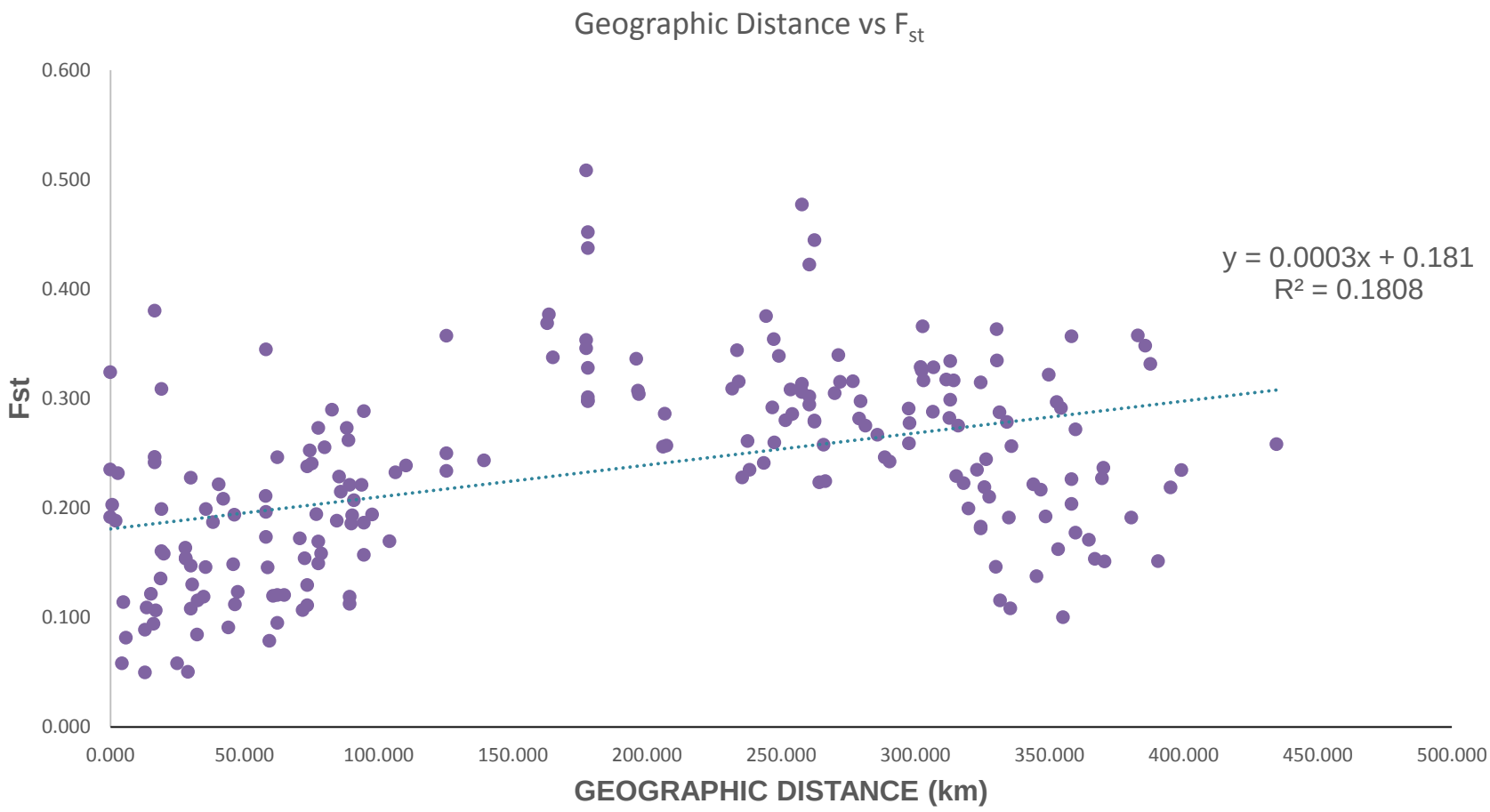


Figure 3. Cumulative Pairwise F_{ST} Values Between Each Tributary and Geographic Location For All Sampling Sites Within California. Performing a mantel test showed a moderate level of correlation between genetic and geographic distance (km). (R^2=0.1808, p<0.001)

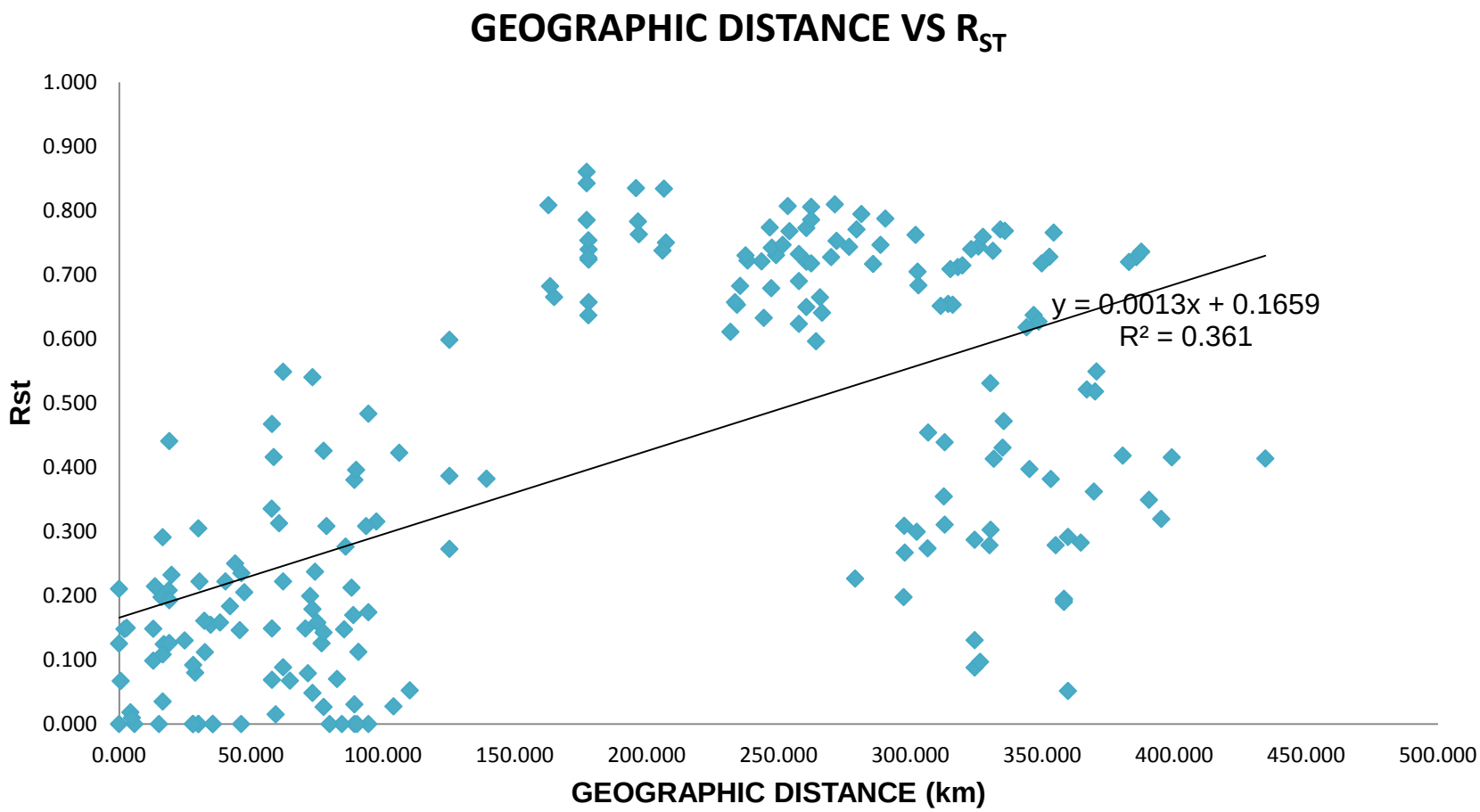


Figure 4. Cumulative Pairwise R_{ST} Values Between Each Tributary and Geographic Location For All Sampling Sites Within California. Performing a mantel test showed a moderate level of correlation between genetic and geographic distance (km) (R^2=0.361, p=0.010)

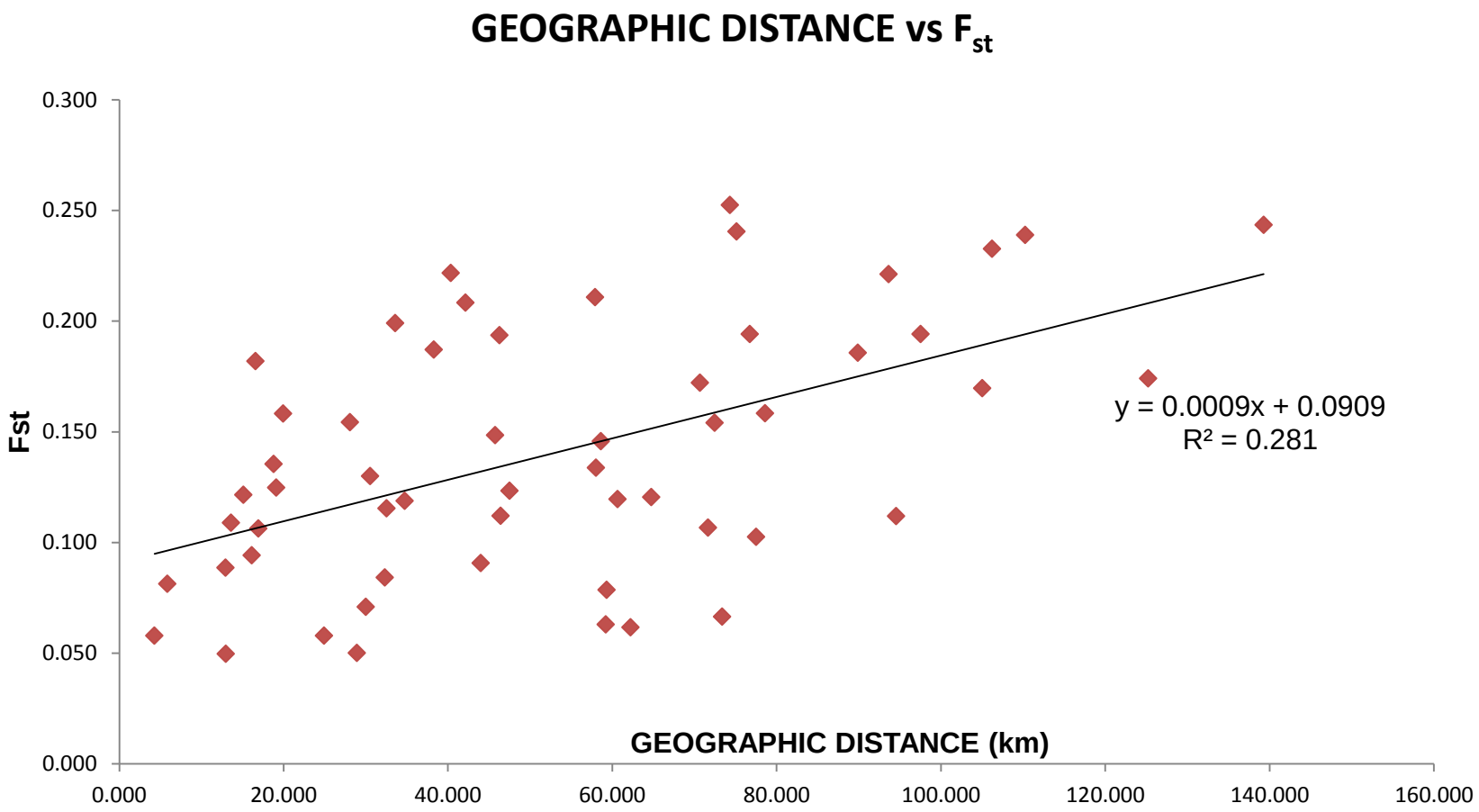


Figure 5. Cumulative Pairwise F_{ST} Values Between Each Tributary and Geographic Location For All Sampling Sites Within the Southern California Region. Performing a mantel test showed a moderate level of correlation between genetic and geographic distance (km).

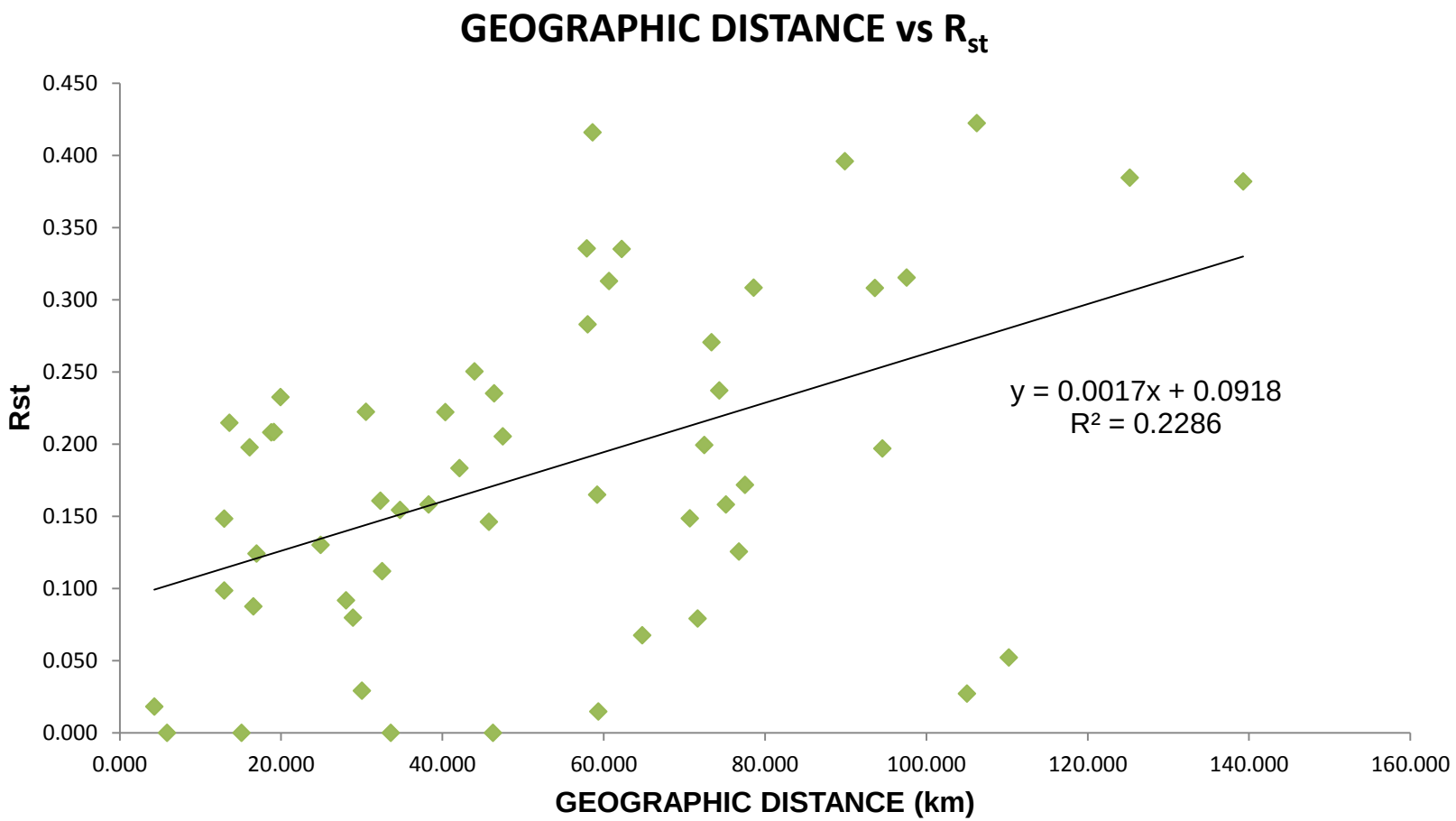


Figure 6. Cumulative Pairwise R_{ST} Values Between Each Tributary and Geographic Location For All Sampling Sites Within California. Performing a mantel test showed a moderate level of correlation between genetic and geographic distance (km) (R^2=0.2286, p=0.010)

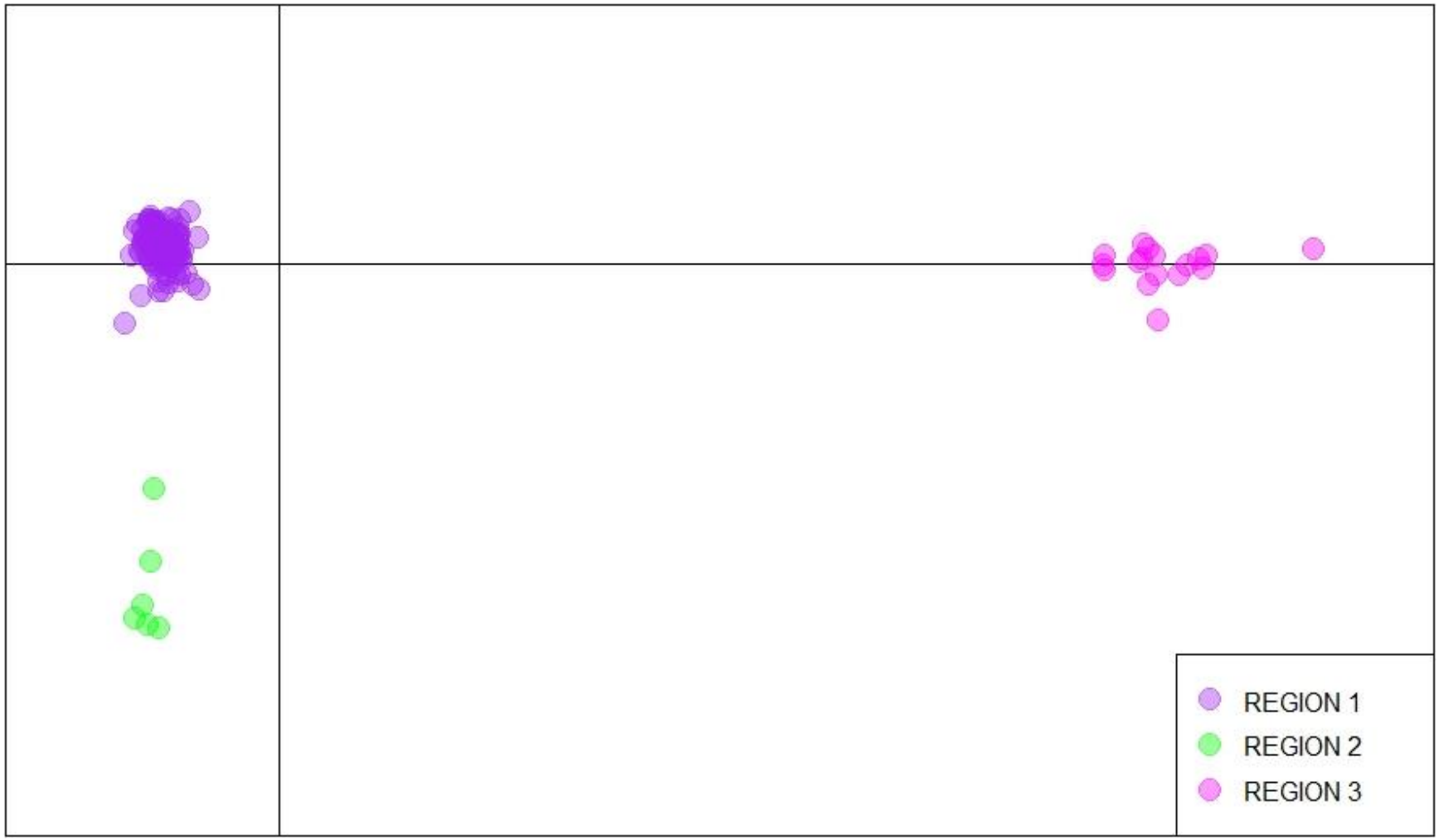


Figure 7. Discriminate Analysis of Principle Components (DAPC) For the First Two Axes Which Identifies K=3. Where 1 represents the southern California populations, 2 represents the Owens River populations; and 3 represents the Central Coast populations.

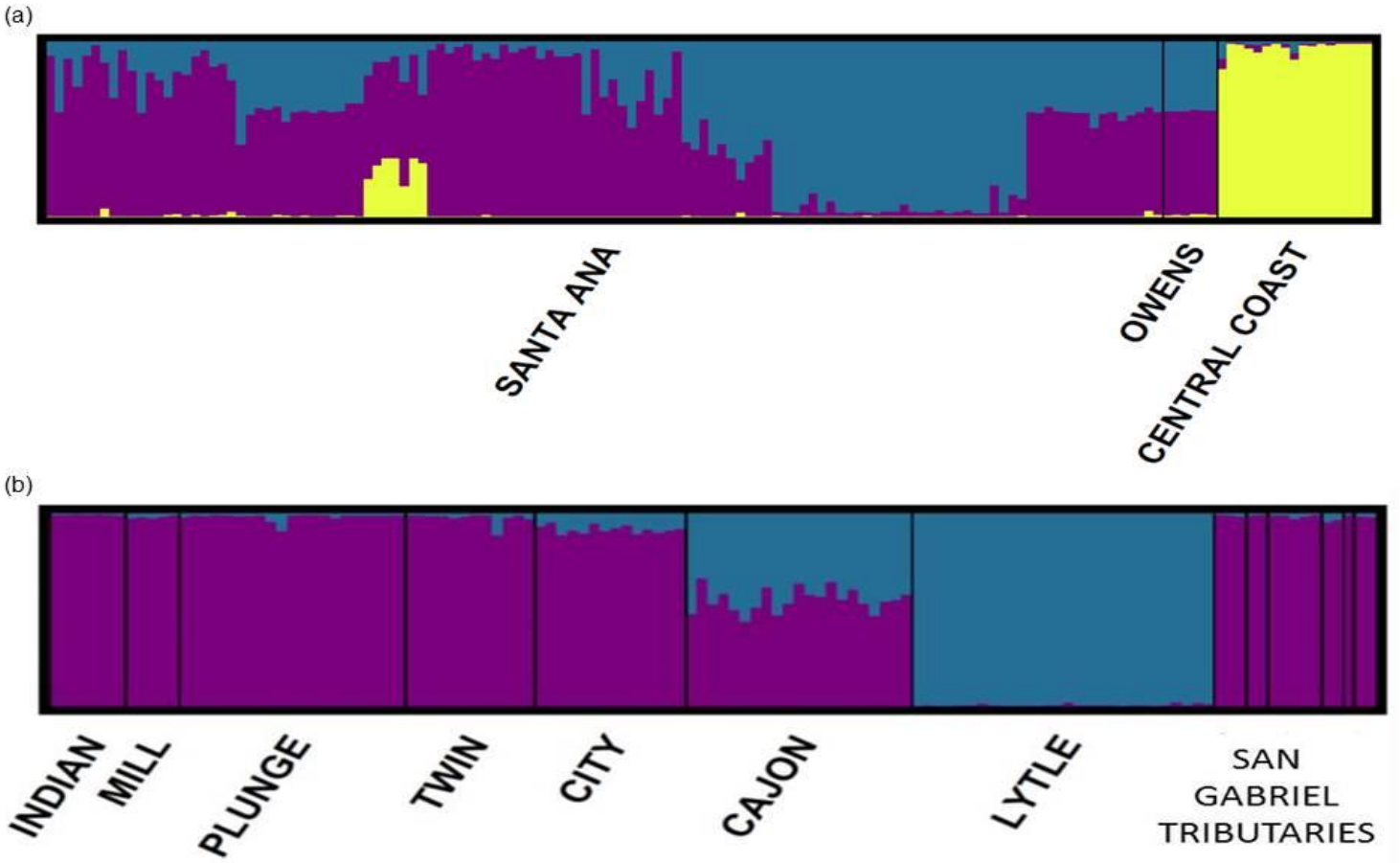


Figure 8. STRUCTURE Results for (a) Three Genetic Structures (K=3) for the Three Regions Sampled in California; Southern California, Owens Valley and Central Coast. Vertical Bars Represent 146 Speckled Dace Samples, While Color Represents the Proportion of Ancestry From Each Population. (b) Two Genetic Clusters (K=2) That Were Identified From the 123 individuals of the Southern California Santa Ana Speckled Dace Populations.

CHAPTER FOUR

DISCUSSION

Microsatellite Analysis

The characterization and identification of the 23 polymorphic microsatellite loci for the Speckled Dace that we established have shown to be successful markers when examining the populations within California. The seven loci that I chose to utilize in this study were able to identify population structure and patterns of gene flow occurring within the Santa Ana Specked Dace but also within the speckled dace populations found in the Central Coast of California and Owens Valley. The numbers of alleles identified per locus (19-26) were comparable to the alleles identified in other microsatellites studies done with the speckled dace (Hoekzema & Sidlauskas, 2014; Kinziger, Nakamoto, Anderson, & Harvey, 2011). In addition, all loci utilized amplified in 100% of the individuals sampled.

Population Structure among the Three Regions

My study revealed a lack of connectivity between the three regions sampled in California. This is some of the first genetic evidence supporting a hypothesis that the Santa Ana Specked Dace populations (Southern California populations) are discontinuous from the other two regions sampled. Current mtDNA studies being performed in Tony Metcalf's lab with *cytochrome b* and *d-loop* (Jay VanMeter, unpublished data) suggest a 6-8% genetic difference

between the Southern California populations and those of the Central Coast and Owens Valley, inferring independent evolutionary trajectories. The microsatellite data corroborates the mtDNA evidence, providing a clear picture of the genetic structure among these regional populations.

Smith & Dowling (2008) suggested that the Los Angeles Basin populations diverged from the Colorado Basin ~1.9 mya. They further suggest that the Owen's Valley populations would have colonized that region sometime between the divergence from a sister species of *R. osculus* ~6.3 mya and the colonization event that occurred in the Colorado Basin, ~3.6 mya. The Owen's Valley dace and the Southern California populations should show more similarity to one another than to those on the Central Coast. Microsatellite genotypes suggest that the Speckled Dace populations differentiate into three genetic populations as is supported with the DAPC (Figure 7). Based on this hypothesis, we would expect that the Owen's Valley dace and the Santa Ana Speckled dace to be more similar to one another, in that they diverged from a more recent common ancestor than the common ancestral populations that would be associated the Central Coast populations. Following the same model, we would expect the Owens Valley dace to show less genetic structure with the Central Coast dace due to less divergence events than those events leading to the Southern California populations (Figure 9). This is clearly supported when examining pairwise R_{ST} values which account for the stepwise mutation model exhibited by microsatellites, rather than the infinite allele model observed with F_{ST} values.

Significant, high levels of differentiation between the regional populations is exhibited (Table 7a, R_{ST} = 0.406-0.659) but greatest between Santa Ana Speckled Dace and those of the Central Coast (R_{ST}=0.659). The AMOVA analysis indicated that 60.0% of the genetic variance observed in the microsatellite loci is due to the difference between each of the regions rather than within each region. Additionally, isolation by distance may be contributing to the structure observed among the sampled regions. This suggests that the gene flow between each of the regions is absent and has been for some time.

It was also hypothesized that the Central Coast populations were established from the migration of the Southern California populations into the region (Cornelius, 1969). The microsatellite data does not support this hypothesis. A high level of population differentiation occurs between the Southern California populations and that of the Central Coast populations (Table 8a: F_{ST}=0.143; Table 7a: R_{ST}=0.659). This would suggest that the Central Coast populations were established from a colonization event that most likely occurred from the population that diverged independently from those that colonized Owens Valley, Colorado Basin, and the Southern California populations. In addition, when examining the membership coefficient of each sampled individual, three inferred populations were identified. The Central Coast populations show minimal admixture with the other two inferred populations further supporting the idea that the central coast populations diverged independently of the populations that colonized the other two regions. Furthermore, the admixture of the two

inferred populations that are observed in the Owens samples supports the conclusion that the Owens Valley populations and the Southern California populations diverged from the same ancestral lineage and the lack of connectivity among watersheds.

The lack of gene flow (Table 6, N_M=0.168) between each of the regions is partially due to geographic distance (Mantel test, R^2=0.361, p=0.010) but also landscape barriers such as mountain ranges. The Central coast populations are west of the Sierra Nevada Mountains and north of the Transverse Range; the Owen's populations lie east of the Sierra Nevada Mountains and north of the Transverse Range; whereas the Southern California populations lie south of the Sierra Nevada Mountains and south of the Transverse Mountain range. Historical gene flow would have been restricted by the elevation gradients and landscape barriers presented by these mountain ranges.

Current gene flow is further restricted within each region due to the fragmented habitats created through anthropogenic effects. Due to the limited sample sizes for both Central Coast and Owens Valley populations, further sampling may provide a better insight into the population structure occurring within each of the regions. Since the geographic distance between each of the sampled regions is great, sampling of intermediate zones may provide further insight into the divergence events that led to the genetic diversity and population

structure exhibited among each of the regions.

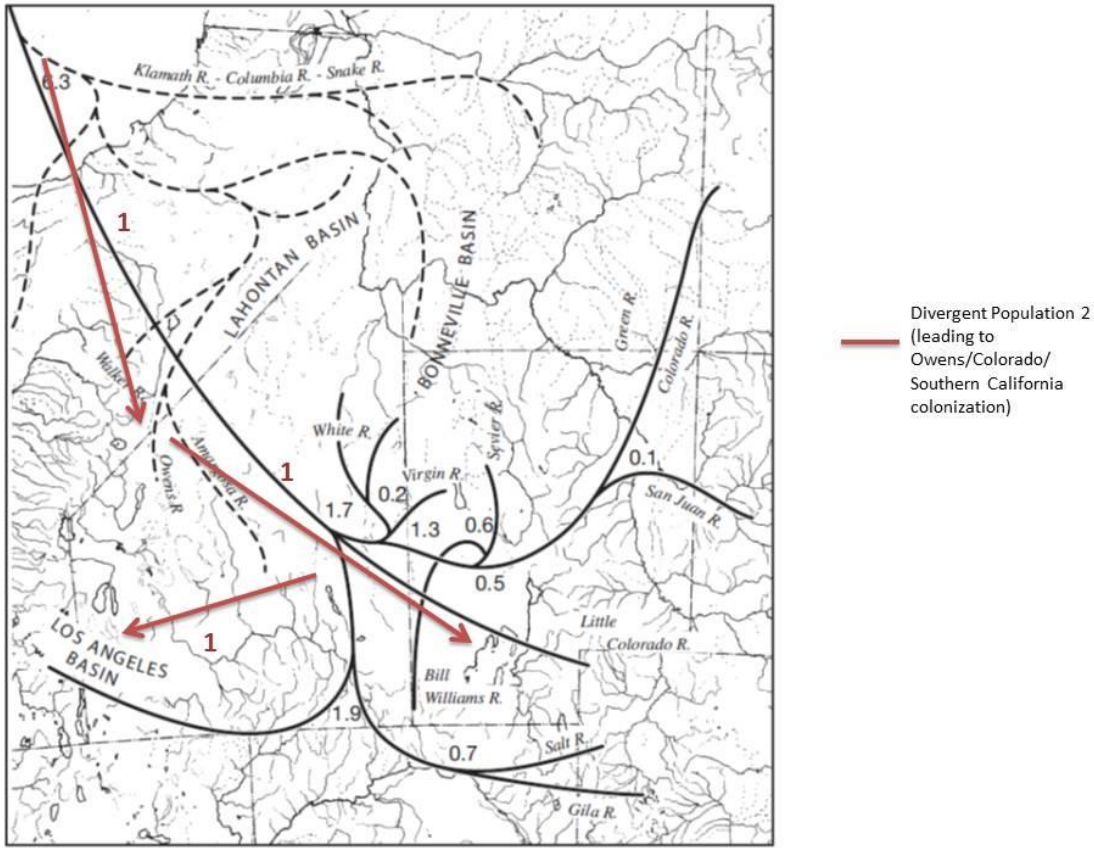


Figure 9. Diagram Depicting Phylogenic Divergence Events Leading to the Colonization of the Santa Ana Speckled Dace Populations of Southern California. Diagram adapted from Smith, G. R., & Dowling, T. E. (2008). Correlating hydrographic events and divergence times of speckled dace (Rhinichthys: Teleostei: Cyprinidae) in the Colorado River drainage. *The Geological Society of America - Special Paper*, 301-315.

Southern California Populations

Overall, my study has revealed that the Santa Ana Speckled Dace populations in Southern California exhibit moderate levels of genetic structure. There is a significant genetic differentiation of the microsatellite genotypes and a

moderate level of population structure correlating to isolation by distance within the Southern California region (Mantel test (F_{ST}/R_{ST})); R^2=.281/R^2=.2286,p=0.01).

This moderate level of population structure was consistent with our expectations due to the highly fragmented habitats that have resulted from both past Pleistocene climate change and more recent anthropogenic effects. During pluvial periods at the end of the Pleistocene, connectivity among the tributaries in the Los Angeles Basin would have been at its peak due to the melting of the Wisconsinian glacier (Colburn, 2006). Up to around a century ago, connectivity still may have been achieved during pluvial events, but then anthropogenic advances led to the restricted flow of most of the tributaries in the region. This would have restricted gene flow between the populations of the Santa Ana Speckled Dace. STRUCTURE analysis indicated that Santa Ana Speckled Dace consists of two ancestral populations (K=2), suggesting that the Southern California populations are not a single panmictic group but rather genetic structure has occurred between the populations.

Four models have been proposed to examine populations structure and gene flow among the Southern California populations of the Santa Ana Speckled Dace; (1) the watershed in which they inhabit (Santa Ana vs San Gabriel River watersheds), (2) the mountain ranges in which the headwaters for each tributary reside (San Jacinto, San Bernardino or San Gabriel Mountains), (3) each tributary is considered its own distinct population, and (4) isolation by distance.

Watershed Model

Watersheds contain the habitats and environmental structures of the tributaries (Frissell, Liss, Warren, & Hurley, 1986). This can isolate organisms to their perspective watersheds and restrict connectivity. Connectivity between the Santa Ana and San Gabriel River watersheds may have occurred historically, but in the last 100 years, flood control has altered many of the original paths in which the tributaries and main rivers flowed (Santa Ana River Watershed Project Authority, 2004). Connectivity between these two watersheds is currently discontinuous due to the fragmented habitats, providing a perfect model to examine current gene flow and population structure. Upon analysis of the microsatellite data utilizing an AMOVA, 15.1% of the genetic variation occurred as a result of differences among the watershed rather than within the watershed, suggesting a moderate level of population structure is occurring between the two watersheds (N_M=1.447).

The San Gabriel watershed's tributaries exhibit very little population structure, providing evidence that the tributaries sampled from this watershed represent a panmictic population (Appendix B, R_{ST}=0.00-0.193). Many of the samples were obtained from locations near the confluences of these tributaries to the San Gabriel River, allowing for possible gene flow between each of the sampled tributaries. Due to the limited sample size obtained from the Hain River (n=2) (Los Angeles River Watershed) and the proximate location to the San Gabriel River, Hain was considered part of the San River Watershed for this

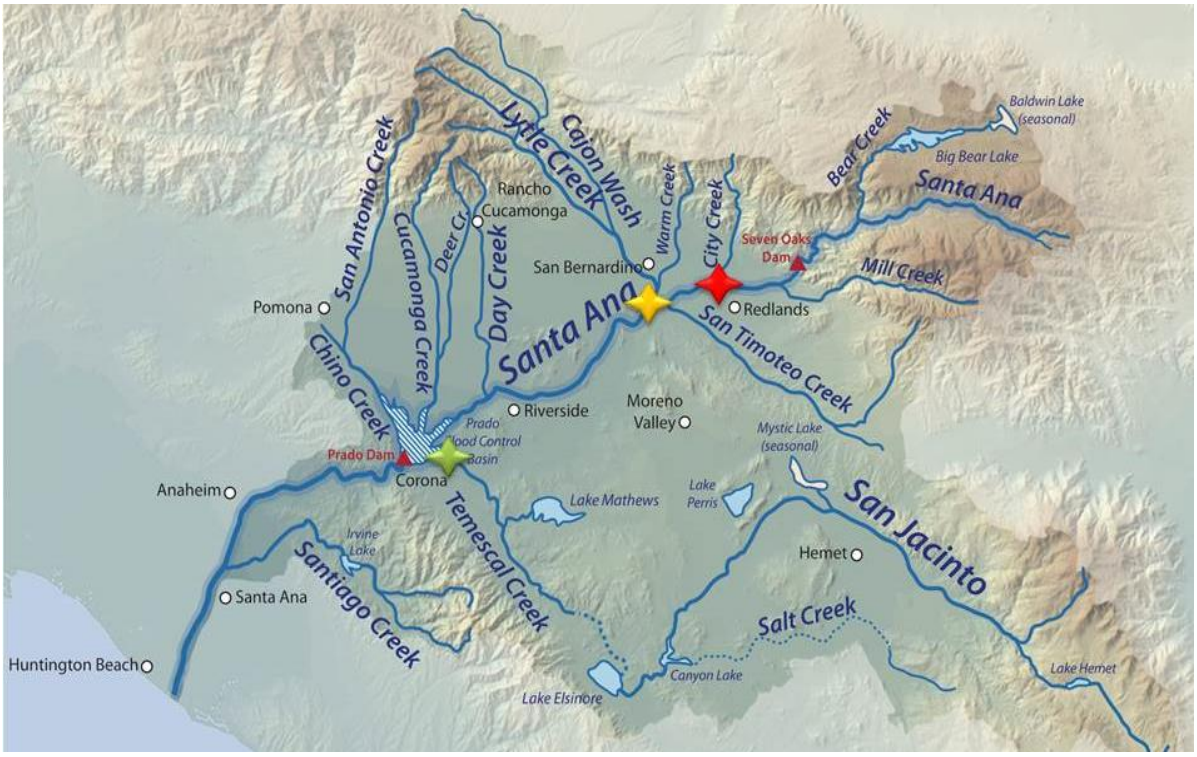


Figure 10. Map Representing the Santa Ana River Watershed Where Colored Stars Represent Confluences of Interest. The red star represents the northeastern most tributaries (Mill, City, Plunge and Twin Creek) confluence; the yellow star represents the Cajon/Lytle confluence; and the green star represents Indian Creek's confluence. Map adapted from (Shannon1. (2015, March 27). *Map of the Santa Ana River basin.* Retrieved from Newkis: http://www.newikis.com/en/commons/File:Santa-ana-river-new.jpg)

study. Very little genetic structure occurred among Hain and the San Gabriel tributaries (R_{ST}=0.00-0.291), indicating that gene flow would have occurred among the Los Angeles and San Gabriel River watershed. The pairwise R_{ST} values for within the San Gabriel tributaries are comparable to those of Hain and the San Gabriel tributaries further supporting this gene flow model.

Mountain Range Model

The Southern California region is comprised of various mountain ranges that could account for different genetic breaks in organisms. For Santa Ana Speckled Dace populations, three mountain ranges contain the headwaters for the tributaries in question. The San Gabriel Mountains provide the headwaters to Mill, Plunge, Twin, and City Creek. The San Jacinto Mountains are the location for the headwaters to Indian Creek. The San Gabriel Mountains contain the headwaters to all forks of the San Gabriel River, including Fish Creek and Cattle Canyon, as well as Lytle and Cajon Creek of the Santa Ana River watershed. Numerous freshwater organism studies have identified the Transverse Range Break (separating the San Gabriel and San Bernardino Mountains) as a distinct phylogenetic lineage break between populations (Phillipsen & Metcalf, 2009; Spinks, Thomson, & Shaffer, 2010; Chatzimanolis & Caterino, 2007). This lineage break is not well supported by the microsatellite data. The AMOVA indicated that very little to no population structure was observed between mountain ranges. 7.7% of the genetic variance was among mountain ranges and the migration rate was highest (N_M=3.134) in this model. This would suggest that the mountains in which the headwaters for each tributary are located do not act as genetic barriers for Santa Ana Speckled Dace. Cajon and Lytle Creek are part of the Santa Ana watershed and may have had historical connectivity with the other tributaries within the Santa Ana watershed. Therefore, the mountain

range model does not accurately reflect the current geographic structure or the patterns of gene flow.

Tributary Model

Historic connectivity between adjacent streams would have been common prior to the establishment of county-controlled flood measures. As a result of these flood control measures and an increase in human population density within the Southern California region, tributaries have become highly fragmented and isolated from each other. Most of the Santa Ana Speckled Dace populations reside in the highlands of each tributary, preventing further connectivity among tributaries. As a result of their increased isolation from each other, tributary populations are rarely sharing migrants. The greatest levels of significant population structure occur under the tributary model, where each tributary is considered its own population (R_{ST}=0.160). The values most likely indicate a degree of historical population structure, but given the current flood control measures instituted by local county and state agencies, population structure is expected to increase.

Isolation by Distance

Analysis of the microsatellite data has revealed a significant correlation between population structure and geographic distance among the tributaries within the Southern California region (R^2=0.2286, p-value=0.010; Figure 6), suggesting isolation by distance. The microsatellite data suggests that the populations found within the tributaries that occupy the northeast region of the

Santa Ana watershed (Mill, Plunge, City and Twin) would have had the greatest opportunity for gene flow during pluvial events. In addition, the confluences for each of the tributaries with the Santa Ana River are very close to one another (Figure 10), allowing further opportunities for gene flow (R_{ST}=0.00-0.222). Lytle and Cajon Creek are located in the northwestern region of the Santa Ana watershed and are proximate to one another. Greater amounts of gene flow would be expected between these two tributaries as is supported by the microsatellite data (Appendix, R_{ST}=.018). The confluence for Lytle/Cajon is found south of the confluence for the northeastern tributaries (Figure 10), resulting in possible gene flow in the most proximate tributaries to the confluence, but should decline with geographic distance from the confluence (Appendix B). Indian Creek's confluence is the southernmost confluence (Figure 10) within the sample for the Santa Ana watershed. Minimal gene flow would be expected between the northernmost populations and Indian Creek, as is evident with the microsatellite data (Appendix B, R_{ST}=0.146-0.335). This east to west pattern is supported by STRUCTURE (Figure 8b) where the contributions of the two ancestral populations stretch proportionally from Indian Creek in east to Lytle Creek in the west.

Conservation Implications

The Santa Ana Speckled Dace has been listed as a Species of Concern by the California Department of Fish and Wildlife and the United State Forest Service. Due to the effects of fires and floods in the area, some of the

populations have more recently become extirpated. The population genetic data provided by this study has important implications in the conservation management strategies for the Santa Ana Speckled Dace. Habitat loss and fragmentation has placed the Santa Ana Speckled Dace on the pathway to possible extinction within the Southern California region. Introduction and reintroduction of the Santa Ana Speckled Dace populations has been attempted by the United States Forest Service within Lytle Creek, but with no recorded success. Since our lab began work on the Santa Ana Speckled Dace, three of the seven tributary populations within the Santa Ana Watershed have become extirpated as a result of fire and floods and may be reintroduced in the future. This study has revealed that moderate levels of population structure occur between each of the tributaries within the Southern California region due to restricted gene flow. Therefore, it is important to consider this when deciding upon a stock population to be utilized for such purposes.

It has been shown that geographic distance and genetic differentiation are moderately correlated with each other. This is important to consider when re-establishing an extirpated population. By choosing a tributary that is in proximal location to the extirpated population, a better representation of the original ancestral alleles that were present in that tributary may be re-established. In addition, the microsatellite data has provided a baseline for the most recent genetic composition of each of the tributaries within the Southern California region. Utilizing this information, we can monitor the changes in the genetic

composition of the populations to further evaluate the effects of habitat fragmentation and isolation on the Santa Ana Speckled Dace populations.

The Santa Ana Speckled Dace populations have shown to be genetically distinct from other regions within California. With further evaluation of the mitochondrial and nuclear genome, being performed by our lab, the Santa Ana Speckled Dace may show to be reciprocally monophyletic for all markers, suggesting that the Southern California population be a distinct taxa, or species.

Conclusions

The speckled dace of the Central Coast, Owens Valley and Southern California have shown to be discontinuous, highly differentiated populations, with no current gene flow occurring between the regions. The data presented in this study supports the findings of Smith & Dowling (2008) suggesting that divergence events led to the Owens Valley, Colorado Basin and Southern California populations. The microsatellite data revealed that the Southern California populations do not share ancestry with the Central Coast populations as was previously hypothesized. Further analysis needs to be performed examining the intermediate ranges between the three geographic regions sampled to provide further insight into the divergent events that led to the colonization of the extant populations.

The Santa Ana Speckled Dace populations within Southern California exhibit moderate levels of population structure and genetic variation. Habitat loss and fragmentation resulting from anthropogenic effects has restricted gene flow

resulting in the greatest levels of genetic differentiation occurring among each tributary within the Southern California region. Current levels of genetic diversity and gene flow have been established for the Santa Ana Speckle Dace. This study shows that the Santa Ana Speckled Dace is minimally a conservation management unit, and has established a genetic profile for each of the tributaries that can be utilized in conservation management strategies to maintain the local biodiversity of the Santa Ana Speckled Dace.

APPENDIX A

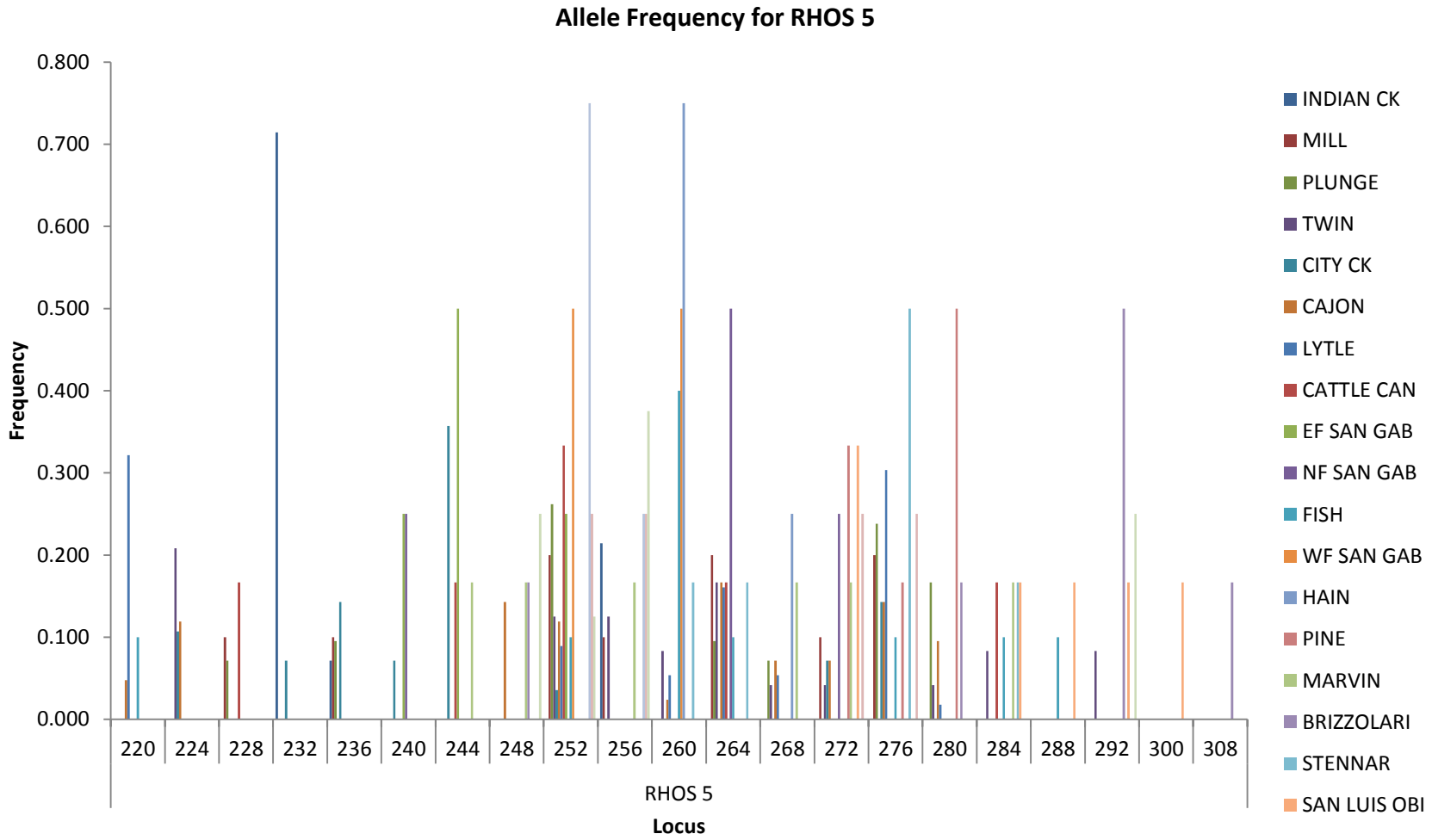
MICROSATELLITE DATA

RHOS5 ALLELE FREQUENCIES (SOUTHERN CALIFORNIA POPULATIONS)

Locus	Allele/n		IND	PLNG	TWIN	CITY	CAJON	LYTLE	CATTLE	EFSGR	NFSGR	FISH	WFSGR	HAIN
RHOS 5	**N**	7	5	21	12	14	21	28	3	2	2	5	1	2
	220	0.000	0.000	0.000	0.000	0.000	0.048	0.321	0.000	0.000	0.000	0.100	0.000	0.000
	224	0.000	0.000	0.000	0.208	0.107	0.119	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	228	0.000	0.100	0.071	0.000	0.000	0.000	0.000	0.167	0.000	0.000	0.000	0.000	0.000
	232	0.714	0.000	0.000	0.000	0.071	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	236	0.071	0.100	0.095	0.000	0.143	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	240	0.000	0.000	0.000	0.000	0.071	0.000	0.000	0.000	0.250	0.250	0.000	0.000	0.000
	244	0.000	0.000	0.000	0.000	0.357	0.000	0.000	0.167	0.500	0.000	0.000	0.000	0.000
	248	0.000	0.000	0.000	0.000	0.000	0.143	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	252	0.000	0.200	0.262	0.125	0.036	0.119	0.089	0.333	0.250	0.000	0.100	0.500	0.000
	256	0.214	0.100	0.000	0.125	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	260	0.000	0.000	0.000	0.083	0.000	0.024	0.054	0.000	0.000	0.000	0.400	0.500	0.750
	264	0.000	0.200	0.095	0.167	0.000	0.167	0.161	0.167	0.000	0.500	0.100	0.000	0.000
	268	0.000	0.000	0.071	0.042	0.000	0.071	0.054	0.000	0.000	0.000	0.000	0.000	0.250
	272	0.000	0.100	0.000	0.042	0.071	0.071	0.000	0.000	0.000	0.250	0.000	0.000	0.000
	276	0.000	0.200	0.238	0.000	0.143	0.143	0.304	0.000	0.000	0.000	0.100	0.000	0.000
	280	0.000	0.000	0.167	0.042	0.000	0.095	0.018	0.000	0.000	0.000	0.000	0.000	0.000
	284	0.000	0.000	0.000	0.083	0.000	0.000	0.000	0.167	0.000	0.000	0.100	0.000	0.000
	288	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.100	0.000	0.000
	292	0.000	0.000	0.000	0.083	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	300	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	308	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

RHOS5 ALLELE FREQUENCIES (CENTRAL AND OWENS POPULATIONS)

Locus	Allele/n	PINE	MARVIN	BRIZZOLARI	STENNER	SAN LUIS OBI	DAVY BROWN	CUYAMA	MANZANA
RHOS 5	**N**	3	3	3	3	3	2	2	4
	220	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	224	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	228	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	232	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	236	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	240	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	244	0.000	0.167	0.000	0.000	0.000	0.000	0.000	0.000
	248	0.000	0.167	0.167	0.000	0.000	0.000	0.000	0.250
	252	0.000	0.000	0.000	0.000	0.000	0.750	0.250	0.125
	256	0.000	0.167	0.000	0.000	0.000	0.250	0.250	0.375
	260	0.000	0.000	0.000	0.167	0.000	0.000	0.000	0.000
	264	0.000	0.000	0.000	0.167	0.000	0.000	0.000	0.000
	268	0.000	0.167	0.000	0.000	0.000	0.000	0.000	0.000
	272	0.333	0.167	0.000	0.000	0.333	0.000	0.250	0.000
	276	0.167	0.000	0.000	0.500	0.000	0.000	0.250	0.000
	280	0.500	0.000	0.167	0.000	0.000	0.000	0.000	0.000
	284	0.000	0.167	0.000	0.167	0.167	0.000	0.000	0.000
	288	0.000	0.000	0.000	0.000	0.167	0.000	0.000	0.000
	292	0.000	0.000	0.500	0.000	0.167	0.000	0.000	0.250
	300	0.000	0.000	0.000	0.000	0.167	0.000	0.000	0.000
	308	0.000	0.000	0.167	0.000	0.000	0.000	0.000	0.000

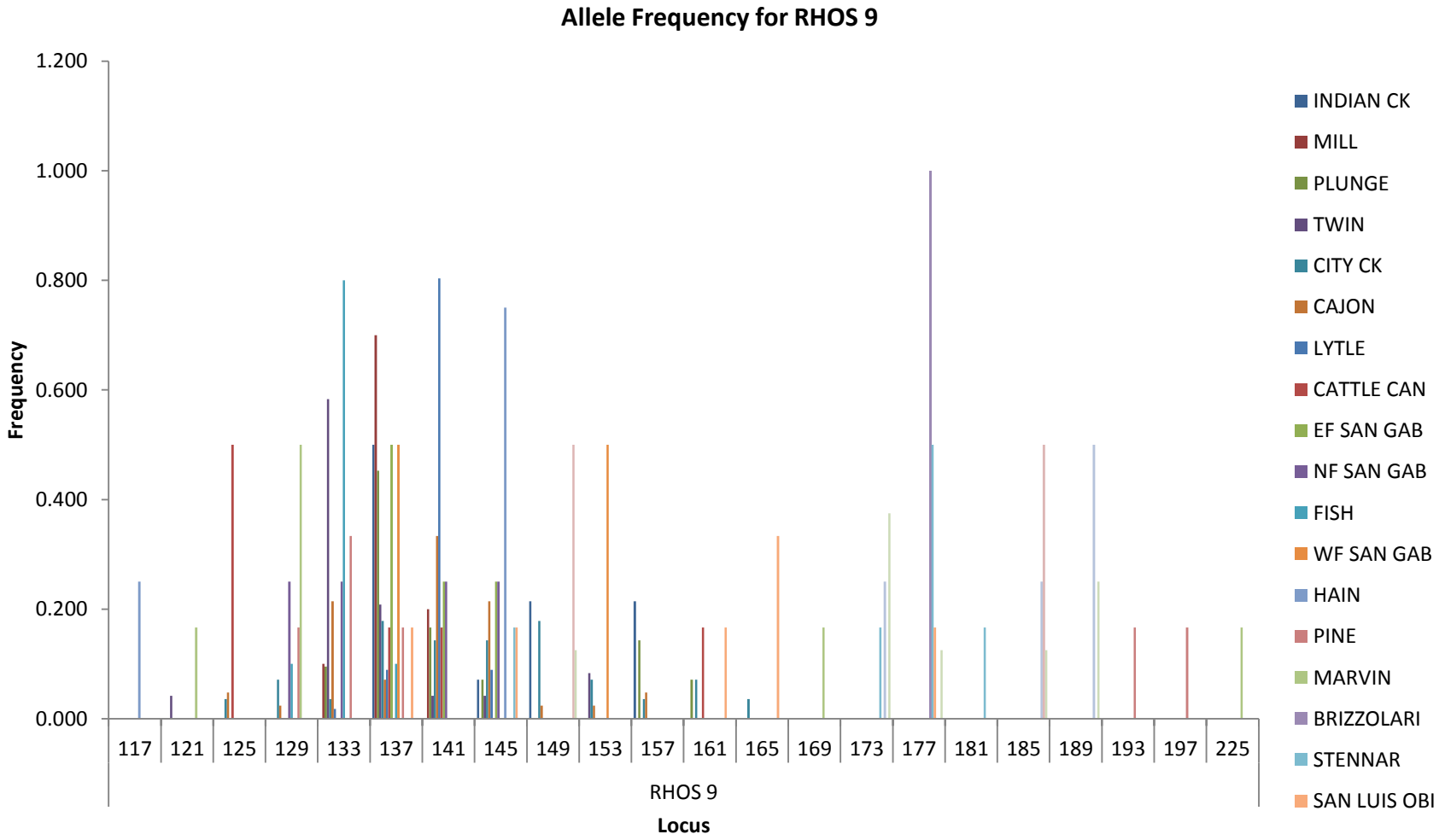

Allele Frequency for RHOS 5
0.800
0.700
0.600
0.500
0.400
0.300
0.200
0.100
0.000
Frequency
220
224
228
232
236
240
244
248
252
256
260
264
268
272
276
280
284
288
292
300
308
RHOS 5
Locus
INDIAN CK
MILL
PLUNGE
TWIN
CITY CK
CAJON
LYTLE
CATTLE CAN
EF SAN GAB
NF SAN GAB
FISH
WF SAN GAB
HAIN
PINE
MARVIN
BRIZZOLARI
STENNAR
SAN LUIS OBI

RHOS9 ALLELE FREQUENCIES (SOUTHERN CALIFORNIA POPULATIONS)

Locus	Allele/n		IND	PLNG	TWIN	CITY	CAJON	LYTLE	CATTLE	EFSGR	NFSGR	FISH	WFSGR	HAIN
RHOS 9	**N**	7	5	21	12	14	21	28	3	2	2	5	1	2
	117	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.250
	121	0.000	0.000	0.000	0.042	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	125	0.000	0.000	0.000	0.000	0.036	0.048	0.000	0.500	0.000	0.000	0.000	0.000	0.000
	129	0.000	0.000	0.000	0.000	0.071	0.024	0.000	0.000	0.000	0.250	0.100	0.000	0.000
	133	0.000	0.100	0.095	0.583	0.036	0.214	0.018	0.000	0.000	0.250	0.800	0.000	0.000
	137	0.500	0.700	0.452	0.208	0.179	0.071	0.089	0.167	0.500	0.000	0.100	0.500	0.000
	141	0.000	0.200	0.167	0.042	0.143	0.333	0.804	0.167	0.250	0.250	0.000	0.000	0.000
	145	0.071	0.000	0.071	0.042	0.143	0.214	0.089	0.000	0.250	0.250	0.000	0.000	0.750
	149	0.214	0.000	0.000	0.000	0.179	0.024	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	153	0.000	0.000	0.000	0.083	0.071	0.024	0.000	0.000	0.000	0.000	0.000	0.500	0.000
	157	0.214	0.000	0.143	0.000	0.036	0.048	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	161	0.000	0.000	0.071	0.000	0.071	0.000	0.000	0.167	0.000	0.000	0.000	0.000	0.000
	165	0.000	0.000	0.000	0.000	0.036	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	169	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	173	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	177	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	181	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	185	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	189	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	193	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	197	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	225	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

RHOS9 ALLELE FREQUENCIES (CENTRAL AND OWENS POPULATIONS)

Locus	Allele/n	PINE	MARVIN	BRIZZOLARI	STENNER	SAN LUIS OBI	DAVY BROWN	CUYAMA	MANZANA
RHOS 9	**N**	3	3	3	3	3	2	2	4
		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	117	0.000	0.167	0.000	0.000	0.000	0.000	0.000	0.000
	121	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	125	0.167	0.500	0.000	0.000	0.000	0.000	0.000	0.000
	129	0.333	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	133	0.167	0.000	0.000	0.000	0.167	0.000	0.000	0.000
	137	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	141	0.000	0.000	0.000	0.167	0.167	0.000	0.000	0.000
	145	0.000	0.000	0.000	0.000	0.000	0.000	0.500	0.125
	149	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	153	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	157	0.000	0.000	0.000	0.000	0.167	0.000	0.000	0.000
	161	0.000	0.000	0.000	0.000	0.333	0.000	0.000	0.000
	165	0.000	0.167	0.000	0.000	0.000	0.000	0.000	0.000
	169	0.000	0.000	0.000	0.167	0.000	0.250	0.000	0.375
	173	0.000	0.000	1.000	0.500	0.167	0.000	0.000	0.125
	177	0.000	0.000	0.000	0.167	0.000	0.000	0.000	0.000
	181	0.000	0.000	0.000	0.000	0.000	0.250	0.500	0.125
	185	0.000	0.000	0.000	0.000	0.000	0.500	0.000	0.250
	189	0.167	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	193	0.167	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	197	0.000	0.167	0.000	0.000	0.000	0.000	0.000	0.000

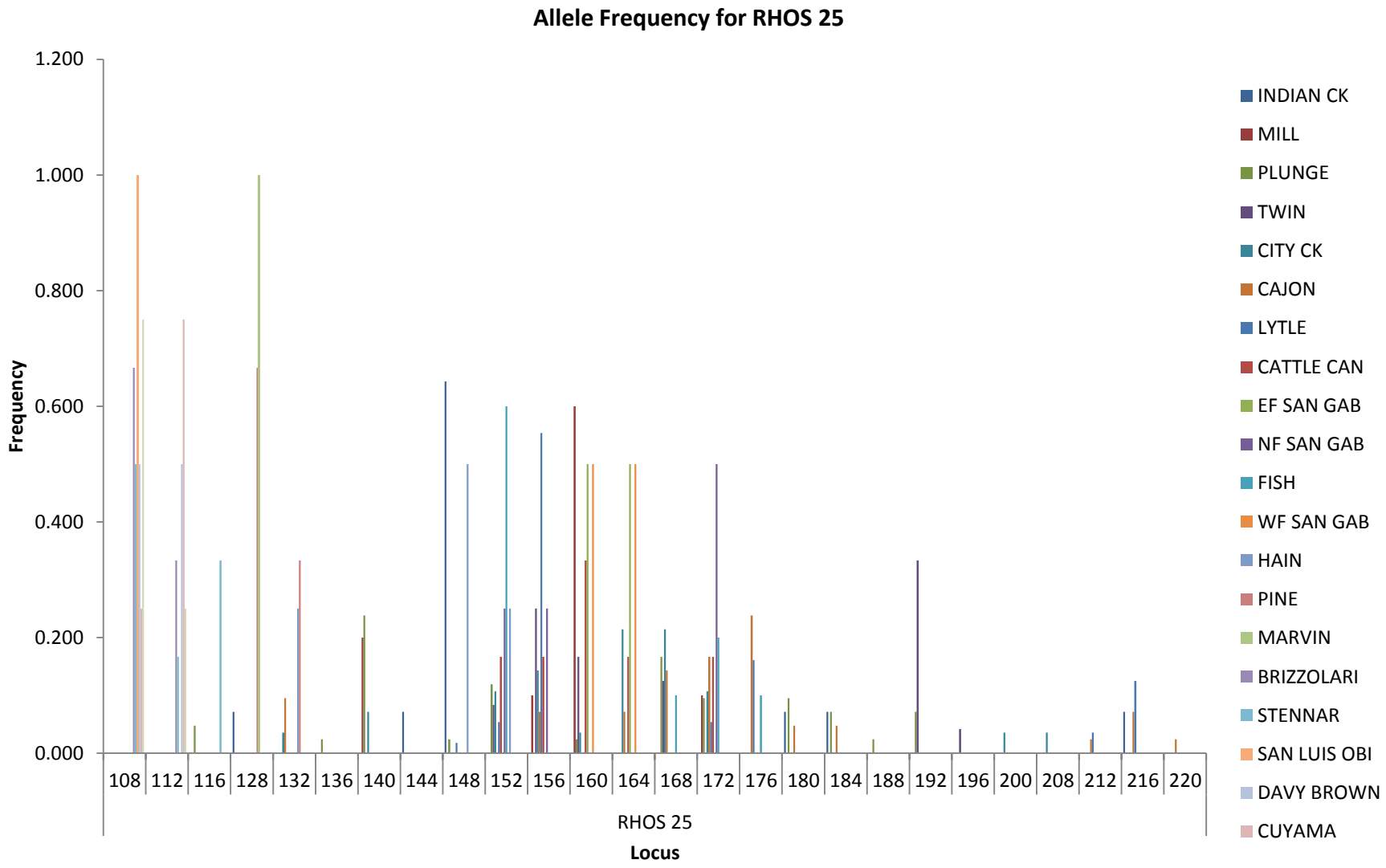

Allele Frequency for RHOS 9
1.200
1.000
0.800
0.600
0.400
0.200
0.000
Frequency
117
121
125
129
133
137
141
145
149
153
157
161
165
169
173
177
181
185
189
193
197
225
RHOS 9
Locus
INDIAN CK
MILL
PLUNGE
TWIN
CITY CK
CAJON
LYTLE
CATTLE CAN
EF SAN GAB
NF SAN GAB
FISH
WF SAN GAB
HAIN
PINE
MARVIN
BRIZZOLARI
STENNAR
SAN LUIS OBI

RHOS25 ALLELE FREQUENCIES (SOUTHERN CALIFORNIA POPULATIONS)

Locus	Allele/n		IND	PLNG	TWIN	CITY	CAJON	LYTLE	CATTLE	EFSGR	NFSGR	FISH	WFSGR	HAIN
RHOS 25	**N**	7	5	21	12	14	21	28	3	2	2	5	1	2
	108	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	112	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	116	0.000	0.000	0.048	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	128	0.071	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	132	0.000	0.000	0.000	0.000	0.036	0.095	0.000	0.000	0.000	0.000	0.000	0.000	0.250
	136	0.000	0.000	0.024	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	140	0.000	0.200	0.238	0.000	0.071	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	144	0.071	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	148	0.643	0.000	0.024	0.000	0.000	0.000	0.018	0.000	0.000	0.000	0.000	0.000	0.500
	152	0.000	0.000	0.119	0.083	0.107	0.000	0.054	0.167	0.000	0.250	0.600	0.000	0.250
	156	0.000	0.100	0.000	0.250	0.143	0.071	0.554	0.167	0.000	0.250	0.000	0.000	0.000
	160	0.000	0.600	0.024	0.167	0.036	0.000	0.000	0.333	0.500	0.000	0.000	0.500	0.000
	164	0.000	0.000	0.000	0.000	0.214	0.071	0.000	0.167	0.500	0.000	0.000	0.500	0.000
	168	0.000	0.000	0.167	0.125	0.214	0.143	0.000	0.000	0.000	0.000	0.100	0.000	0.000
	172	0.000	0.100	0.095	0.000	0.107	0.167	0.054	0.167	0.000	0.500	0.200	0.000	0.000
	176	0.000	0.000	0.000	0.000	0.000	0.238	0.161	0.000	0.000	0.000	0.100	0.000	0.000
	180	0.071	0.000	0.095	0.000	0.000	0.048	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	184	0.071	0.000	0.071	0.000	0.000	0.048	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	188	0.000	0.000	0.024	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	192	0.000	0.000	0.071	0.333	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	196	0.000	0.000	0.000	0.042	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	200	0.000	0.000	0.000	0.000	0.036	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	208	0.000	0.000	0.000	0.000	0.036	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	212	0.000	0.000	0.000	0.000	0.000	0.024	0.036	0.000	0.000	0.000	0.000	0.000	0.000
	216	0.071	0.000	0.000	0.000	0.000	0.071	0.125	0.000	0.000	0.000	0.000	0.000	0.000
	220	0.000	0.000	0.000	0.000	0.000	0.024	0.000	0.000	0.000	0.000	0.000	0.000	0.000

RHOS25 ALLELE FREQUENCIES (CENTRAL AND OWENS POPULATIONS)

Locus	Allele/n	PINE	MARVIN	BRIZZOLARI	STENNER	SAN LUIS OBI	DAVY BROWN	CUYAMA	MANZANA
RHOS 25	N	3	3	3	3	3	2	2	4
	108	0.000	0.000	0.667	0.500	1.000	0.500	0.250	0.750
	112	0.000	0.000	0.333	0.167	0.000	0.500	0.750	0.250
	116	0.000	0.000	0.000	0.333	0.000	0.000	0.000	0.000
	128	0.667	1.000	0.000	0.000	0.000	0.000	0.000	0.000
	132	0.333	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	136	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	140	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	144	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	148	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	152	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	156	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	160	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	164	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	168	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	172	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	176	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	180	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	184	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	188	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	192	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	196	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	200	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	208	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	212	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	216	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	220	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000


Allele Frequency for RHOS 25
1.200
1.000
0.800
0.600
0.400
0.200
0.000
Frequency
108 112 116 128 132 136 140 144 148 152 156 160 164 168 172 176 180 184 188 192 196 200 208 212 216 220
RHOS 25
Locus
INDIAN CK
MILL
PLUNGE
TWIN
CITY CK
CAJON
LYTLE
CATTLE CAN
EF SAN GAB
NF SAN GAB
FISH
WF SAN GAB
HAIN
PINE
MARVIN
BRIZZOLARI
STENNAR
SAN LUIS OBI
DAVY BROWN
CUYAMA

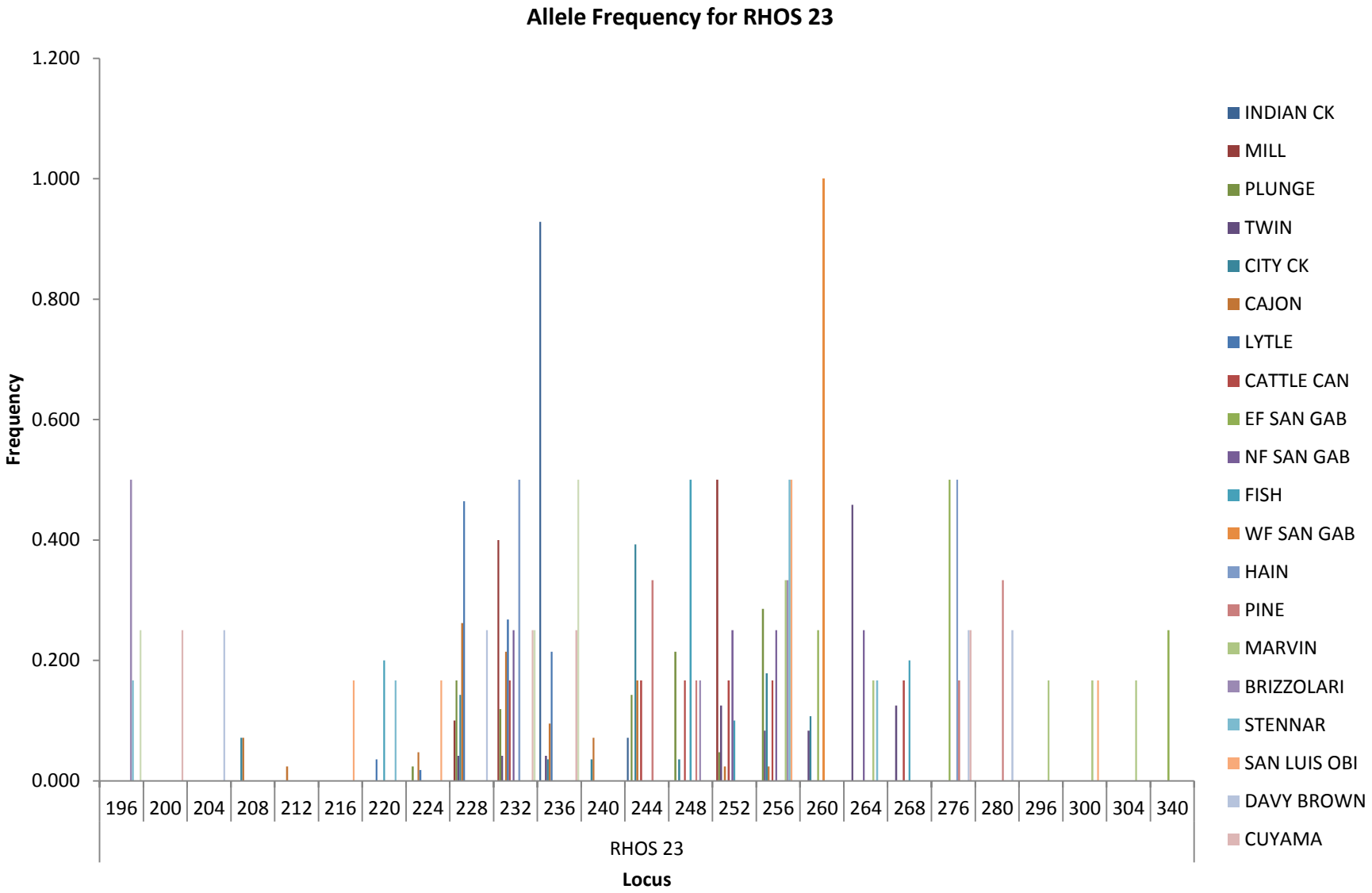
06

RHOS23 ALLELE FREQUENCIES (SOUTHERN CALIFORNIA POPULATIONS)

Locus	Allele/n		IND	PLNG	TWIN	CITY	CAJON	LYTLE	CATTLE	EFSGR	NFSGR	FISH	WFSGR	HAIN
RHOS 23	**N**	7	5	21	12	14	21	28	3	2	2	5	1	2
	196	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	200	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	204	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	208	0.000	0.000	0.000	0.000	0.071	0.071	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	212	0.000	0.000	0.000	0.000	0.000	0.024	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	216	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	220	0.000	0.000	0.000	0.000	0.000	0.000	0.036	0.000	0.000	0.000	0.200	0.000	0.000
	224	0.000	0.000	0.024	0.000	0.000	0.048	0.018	0.000	0.000	0.000	0.000	0.000	0.000
	228	0.000	0.100	0.167	0.042	0.143	0.262	0.464	0.000	0.000	0.000	0.000	0.000	0.000
	232	0.000	0.400	0.119	0.042	0.000	0.214	0.268	0.167	0.000	0.250	0.000	0.000	0.500
	236	0.929	0.000	0.000	0.042	0.036	0.095	0.214	0.000	0.000	0.000	0.000	0.000	0.000
	240	0.000	0.000	0.000	0.000	0.036	0.071	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	244	0.071	0.000	0.143	0.000	0.393	0.167	0.000	0.167	0.000	0.000	0.000	0.000	0.000
	248	0.000	0.000	0.214	0.000	0.036	0.000	0.000	0.167	0.000	0.000	0.500	0.000	0.000
	252	0.000	0.500	0.048	0.125	0.000	0.024	0.000	0.167	0.000	0.250	0.100	0.000	0.000
	256	0.000	0.000	0.286	0.083	0.179	0.024	0.000	0.167	0.000	0.250	0.000	0.000	0.000
	260	0.000	0.000	0.000	0.083	0.107	0.000	0.000	0.000	0.250	0.000	0.000	1.000	0.000
	264	0.000	0.000	0.000	0.458	0.000	0.000	0.000	0.000	0.000	0.250	0.000	0.000	0.000
	268	0.000	0.000	0.000	0.125	0.000	0.000	0.000	0.167	0.000	0.000	0.200	0.000	0.000
	276	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.500	0.000	0.000	0.000	0.500
	280	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	296	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	300	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	304	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	340	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.250	0.000	0.000	0.000	0.000

RHOS23 ALLELE FREQUENCIES (CENTRAL AND OWENS POPULATIONS)

Locus	Allele/n	PINE	MARVIN	BRIZZOLARI	STENNER	SAN LUIS OBI	DAVY BROWN	CUYAMA	MANZANA
RHOS 23	N	3	3	3	3	3	2	2	4
	196	0.000	0.000	0.500	0.167	0.000	0.000	0.000	0.250
	200	0.000	0.000	0.000	0.000	0.000	0.000	0.250	0.000
	204	0.000	0.000	0.000	0.000	0.000	0.250	0.000	0.000
	208	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	212	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	216	0.000	0.000	0.000	0.000	0.167	0.000	0.000	0.000
	220	0.000	0.000	0.000	0.167	0.000	0.000	0.000	0.000
	224	0.000	0.000	0.000	0.000	0.167	0.000	0.000	0.000
	228	0.000	0.000	0.000	0.000	0.000	0.250	0.000	0.000
	232	0.000	0.000	0.000	0.000	0.000	0.000	0.250	0.250
	236	0.000	0.000	0.000	0.000	0.000	0.000	0.250	0.500
	240	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	244	0.333	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	248	0.167	0.000	0.167	0.000	0.000	0.000	0.000	0.000
	252	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	256	0.000	0.333	0.333	0.500	0.500	0.000	0.000	0.000
	260	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	264	0.000	0.167	0.000	0.167	0.000	0.000	0.000	0.000
	268	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	276	0.167	0.000	0.000	0.000	0.000	0.250	0.250	0.000
	280	0.333	0.000	0.000	0.000	0.000	0.250	0.000	0.000
	296	0.000	0.167	0.000	0.000	0.000	0.000	0.000	0.000
	300	0.000	0.167	0.000	0.000	0.167	0.000	0.000	0.000
	304	0.000	0.167	0.000	0.000	0.000	0.000	0.000	0.000
	340	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

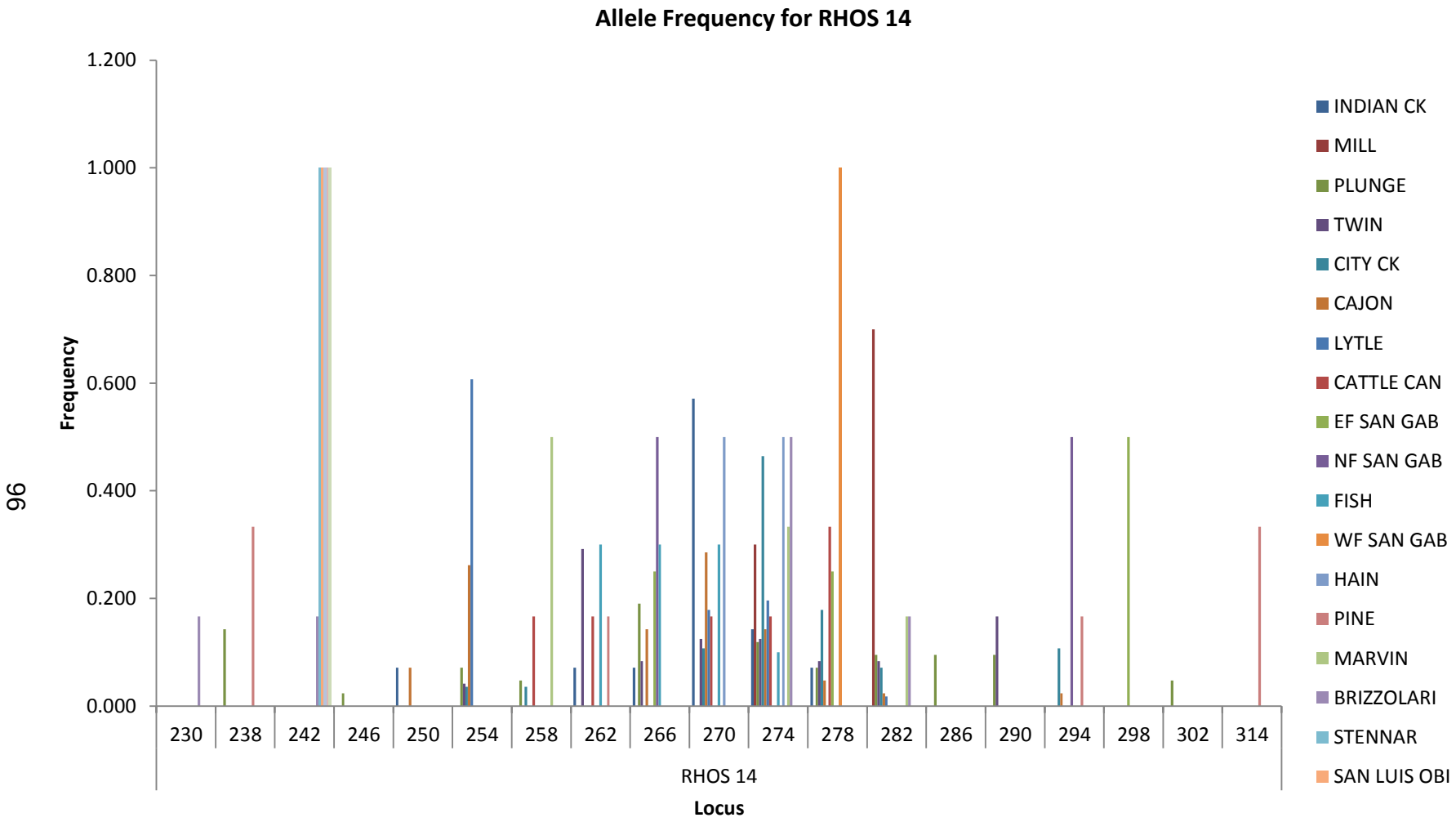

Allele Frequency for RHOS 23
1.200
1.000
0.800
0.600
0.400
0.200
0.000
Frequency
196
200
204
208
212
216
220
224
228
232
236
240
244
248
252
256
260
264
268
276
280
296
300
304
340
RHOS 23
Locus
INDIAN CK
MILL
PLUNGE
TWIN
CITY CK
CAJON
LYTLE
CATTLE CAN
EF SAN GAB
NF SAN GAB
FISH
WF SAN GAB
HAIN
PINE
MARVIN
BRIZZOLARI
STENNAR
SAN LUIS OBI
DAVY BROWN
CUYAMA

RHOS14 ALLELE FREQUENCIES (SOUTHERN CALIFORNIA POPULATIONS)

LOCUS	Allele/n		IND	PLNG	TWIN	CITY	CAJON	LYTLE	CATTLE	EFSGR	NFSGR	FISH	WFSGR	HAIN
RHOS 14	**N**	7	5	21	12	14	21	28	3	2	2	5	1	2
	230	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	238	0.000	0.000	0.143	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	242	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	246	0.000	0.000	0.024	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	250	0.071	0.000	0.000	0.000	0.000	0.071	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	254	0.000	0.000	0.071	0.042	0.036	0.262	0.607	0.000	0.000	0.000	0.000	0.000	0.000
	258	0.000	0.000	0.048	0.000	0.036	0.000	0.000	0.167	0.000	0.000	0.000	0.000	0.000
	262	0.071	0.000	0.000	0.292	0.000	0.000	0.000	0.167	0.000	0.000	0.300	0.000	0.000
	266	0.071	0.000	0.190	0.083	0.000	0.143	0.000	0.000	0.250	0.500	0.300	0.000	0.000
	270	0.571	0.000	0.000	0.125	0.107	0.286	0.179	0.167	0.000	0.000	0.300	0.000	0.500
	274	0.143	0.300	0.119	0.125	0.464	0.143	0.196	0.167	0.000	0.000	0.100	0.000	0.500
	278	0.071	0.000	0.071	0.083	0.179	0.048	0.000	0.333	0.250	0.000	0.000	1.000	0.000
	282	0.000	0.700	0.095	0.083	0.071	0.024	0.018	0.000	0.000	0.000	0.000	0.000	0.000
	286	0.000	0.000	0.095	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	290	0.000	0.000	0.095	0.167	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	294	0.000	0.000	0.000	0.000	0.107	0.024	0.000	0.000	0.000	0.500	0.000	0.000	0.000
	298	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.500	0.000	0.000	0.000	0.000
	302	0.000	0.000	0.048	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	314	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

RHOS14 ALLELE FREQUENCIES (CENTRAL AND OWENS POPULATIONS)

Locus	Allele/n	PINE	MARVIN	BRIZZOLARI	STENNER	SAN LUIS OBI	DAVY BROWN	CUYAMA	MANZANA
RHOS 14	**N**	3	3	3	3	3	2	2	4
	230	0.000	0.000	0.167	0.000	0.000	0.000	0.000	0.000
	238	0.333	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	242	0.000	0.000	0.167	1.000	1.000	1.000	1.000	1.000
	246	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	250	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	254	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	258	0.000	0.500	0.000	0.000	0.000	0.000	0.000	0.000
	262	0.167	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	266	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	270	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	274	0.000	0.333	0.500	0.000	0.000	0.000	0.000	0.000
	278	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	282	0.000	0.167	0.167	0.000	0.000	0.000	0.000	0.000
	286	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	290	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	294	0.167	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	298	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	302	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	314	0.333	0.000	0.000	0.000	0.000	0.000	0.000	0.000

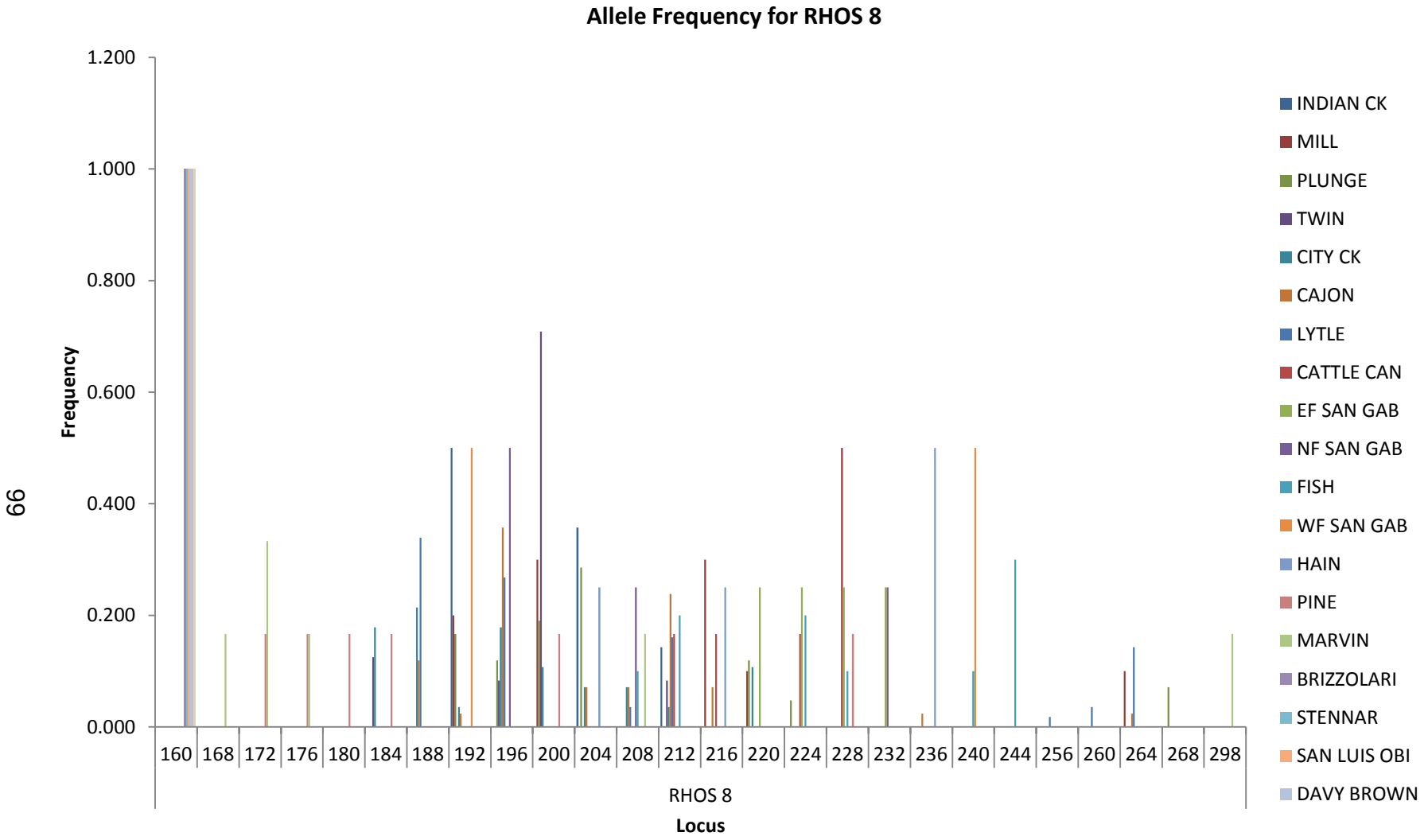

Allele Frequency for RHOS 14
1.200
1.000
0.800
0.600
0.400
0.200
0.000
Frequency
230
238
242
246
250
254
258
262
266
270
274
278
282
286
290
294
298
302
314
RHOS 14
Locus
INDIAN CK
MILL
PLUNGE
TWIN
CITY CK
CAJON
LYTLE
CATTLE CAN
EF SAN GAB
NF SAN GAB
FISH
WF SAN GAB
HAIN
PINE
MARVIN
BRIZZOLARI
STENNAR
SAN LUIS OBI

RHOS8 ALLELE FREQUENCIES (SOUTHERN CALIFORNIA POPULATIONS)

LOCUS	Allele/n		IND	PLNG	TWIN	CITY	CAJON	LYTLE	CATTLE	EFSGR	NFSGR	FISH	WFSGR	HAIN
RHOS 8	**N**	7	5	21	12	14	21	28	3	2	2	5	1	2
	160	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	168	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	172	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	176	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	180	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	184	0.000	0.000	0.000	0.125	0.179	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	188	0.000	0.000	0.000	0.000	0.214	0.119	0.339	0.000	0.000	0.000	0.000	0.000	0.000
	192	0.500	0.200	0.167	0.000	0.036	0.024	0.000	0.000	0.000	0.000	0.000	0.500	0.000
	196	0.000	0.000	0.119	0.083	0.179	0.357	0.268	0.000	0.000	0.500	0.000	0.000	0.000
	200	0.000	0.300	0.190	0.708	0.107	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	204	0.357	0.000	0.286	0.000	0.071	0.071	0.000	0.000	0.000	0.000	0.000	0.000	0.250
	208	0.000	0.000	0.000	0.000	0.071	0.071	0.036	0.000	0.000	0.250	0.100	0.000	0.000
	212	0.143	0.000	0.000	0.083	0.036	0.238	0.161	0.167	0.000	0.000	0.200	0.000	0.000
	216	0.000	0.300	0.000	0.000	0.000	0.071	0.000	0.167	0.000	0.000	0.000	0.000	0.250
	220	0.000	0.100	0.119	0.000	0.107	0.000	0.000	0.000	0.250	0.000	0.000	0.000	0.000
	224	0.000	0.000	0.048	0.000	0.000	0.000	0.000	0.167	0.250	0.000	0.200	0.000	0.000
	228	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.500	0.250	0.000	0.100	0.000	0.000
	232	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.250	0.250	0.000	0.000	0.000
	236	0.000	0.000	0.000	0.000	0.000	0.024	0.000	0.000	0.000	0.000	0.000	0.000	0.500
	240	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.100	0.500	0.000
	244	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.300	0.000	0.000
	256	0.000	0.000	0.000	0.000	0.000	0.000	0.018	0.000	0.000	0.000	0.000	0.000	0.000
	260	0.000	0.000	0.000	0.000	0.000	0.000	0.036	0.000	0.000	0.000	0.000	0.000	0.000
	264	0.000	0.100	0.000	0.000	0.000	0.024	0.143	0.000	0.000	0.000	0.000	0.000	0.000
	268	0.000	0.000	0.071	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	298	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

RHOS8 ALLELE FREQUENCIES (CENTRAL AND OWENS POPULATIONS)

Locus	Allele/n	PINE	MARVIN	BRIZZOLARI	STENNER	SAN LUIS OBI	DAVY BROWN	CUYAMA	MANZANA
RHOS 8	**N**	3	3	3	3	3	2	2	4
	160	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000
	168	0.000	0.167	0.000	0.000	0.000	0.000	0.000	0.000
	172	0.167	0.333	0.000	0.000	0.000	0.000	0.000	0.000
	176	0.167	0.167	0.000	0.000	0.000	0.000	0.000	0.000
	180	0.167	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	184	0.167	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	188	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	192	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	196	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	200	0.167	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	204	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	208	0.000	0.167	0.000	0.000	0.000	0.000	0.000	0.000
	212	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	216	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	220	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	224	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	228	0.167	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	232	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	236	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	240	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	244	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	256	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	260	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	264	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	268	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	298	0.000	0.167	0.000	0.000	0.000	0.000	0.000	0.000

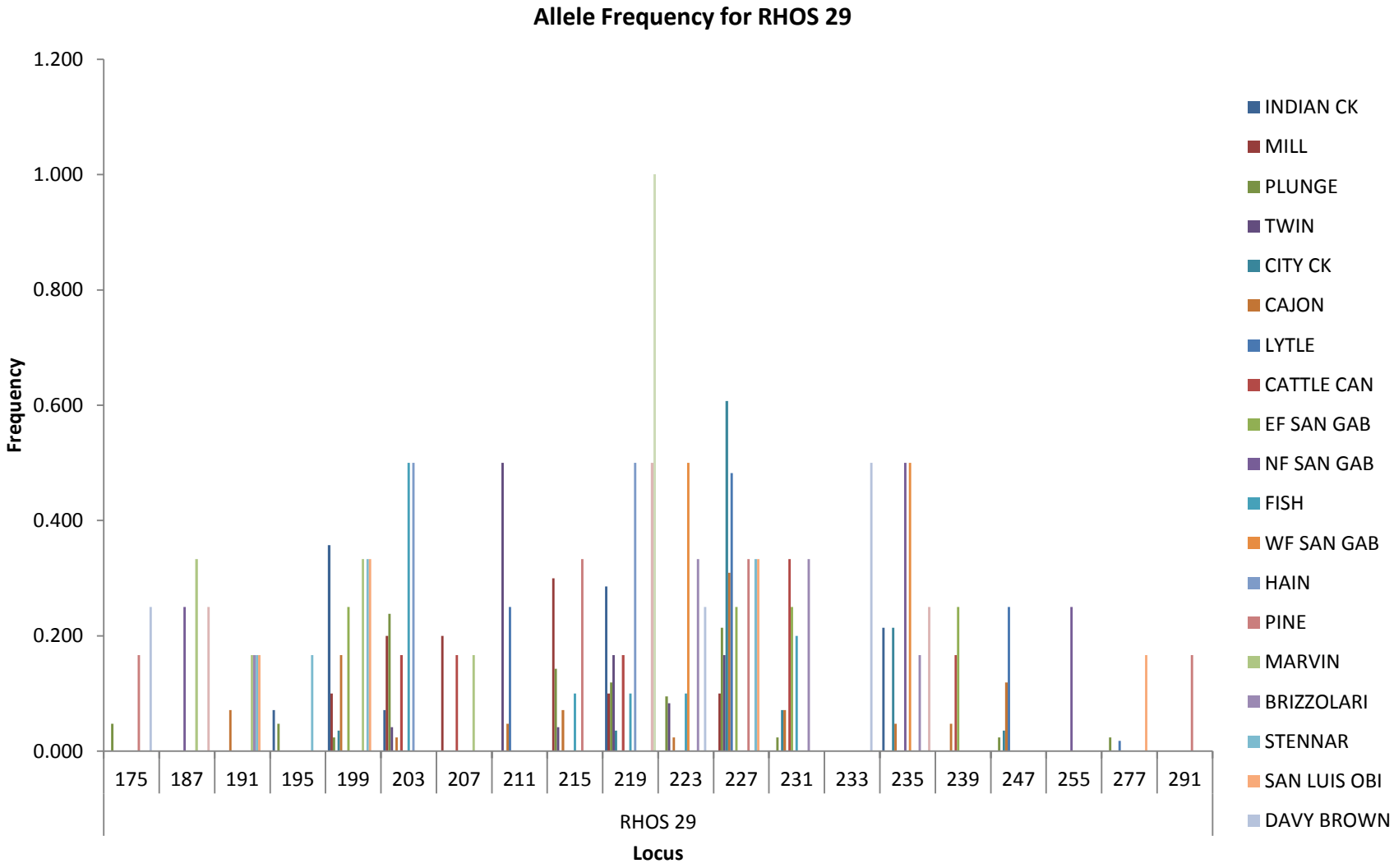

Allele Frequency for RHOS 8
1.200
1.000
0.800
0.600
0.400
0.200
0.000
Frequency
160 168 172 176 180 184 188 192 196 200 204 208 212 216 220 224 228 232 236 240 244 256 260 264 268 298
RHOS 8
Locus
INDIAN CK
MILL
PLUNGE
TWIN
CITY CK
CAJON
LYTLE
CATTLE CAN
EF SAN GAB
NF SAN GAB
FISH
WF SAN GAB
HAIN
PINE
MARVIN
BRIZZOLARI
STENNAR
SAN LUIS OBI
DAVY BROWN

RHOS29 ALLELE FREQUENCIES (SOUTHERN CALIFORNIA POPULATIONS)

LOCUS	Allele/n		IND	PLNG	TWIN	CITY	CAJON	LYTLE	CATTLE	EFSGR	NFSGR	FISH	WFSGR	HAIN
RHOS 29	**N**	7	5	21	12	14	21	28	3	2	2	5	1	2
	175	0.000	0.000	0.048	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	187	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.250	0.000	0.000	0.000
	191	0.000	0.000	0.000	0.000	0.000	0.071	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	195	0.071	0.000	0.048	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	199	0.357	0.100	0.024	0.000	0.036	0.167	0.000	0.000	0.250	0.000	0.000	0.000	0.000
	203	0.071	0.200	0.238	0.042	0.000	0.024	0.000	0.167	0.000	0.000	0.500	0.000	0.500
	207	0.000	0.200	0.000	0.000	0.000	0.000	0.000	0.167	0.000	0.000	0.000	0.000	0.000
	211	0.000	0.000	0.000	0.500	0.000	0.048	0.250	0.000	0.000	0.000	0.000	0.000	0.000
	215	0.000	0.300	0.143	0.042	0.000	0.071	0.000	0.000	0.000	0.000	0.100	0.000	0.000
	219	0.286	0.100	0.119	0.167	0.036	0.000	0.000	0.167	0.000	0.000	0.100	0.000	0.500
	223	0.000	0.000	0.095	0.083	0.000	0.024	0.000	0.000	0.000	0.000	0.100	0.500	0.000
	227	0.000	0.100	0.214	0.167	0.607	0.310	0.482	0.000	0.250	0.000	0.000	0.000	0.000
	231	0.000	0.000	0.024	0.000	0.071	0.071	0.000	0.333	0.250	0.000	0.200	0.000	0.000
	233	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	235	0.214	0.000	0.000	0.000	0.214	0.048	0.000	0.000	0.000	0.500	0.000	0.500	0.000
	239	0.000	0.000	0.000	0.000	0.000	0.048	0.000	0.167	0.250	0.000	0.000	0.000	0.000
	247	0.000	0.000	0.024	0.000	0.036	0.119	0.250	0.000	0.000	0.000	0.000	0.000	0.000
	255	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.250	0.000	0.000	0.000
	277	0.000	0.000	0.024	0.000	0.000	0.000	0.018	0.000	0.000	0.000	0.000	0.000	0.000
	291	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

RHOS29 ALLELE FREQUENCIES (CENTRAL AND OWENS POPULATIONS)

Locus	Allele/n	PINE	MARVIN	BRIZZOLARI	STENNER	SAN LUIS OBI	DAVY BROWN	CUYAMA	MANZANA
RHOS 29	**N**	3	3	3	3	3	2	2	4
	175	0.167	0.000	0.000	0.000	0.000	0.250	0.000	0.000
	187	0.000	0.333	0.000	0.000	0.000	0.000	0.250	0.000
	191	0.000	0.167	0.167	0.167	0.167	0.000	0.000	0.000
	195	0.000	0.000	0.000	0.167	0.000	0.000	0.000	0.000
	199	0.000	0.333	0.000	0.333	0.333	0.000	0.000	0.000
	203	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	207	0.000	0.167	0.000	0.000	0.000	0.000	0.000	0.000
	211	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	215	0.333	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	219	0.000	0.000	0.000	0.000	0.000	0.000	0.500	1.000
	223	0.000	0.000	0.333	0.000	0.000	0.250	0.000	0.000
	227	0.333	0.000	0.000	0.333	0.333	0.000	0.000	0.000
	231	0.000	0.000	0.333	0.000	0.000	0.000	0.000	0.000
	233	0.000	0.000	0.000	0.000	0.000	0.500	0.000	0.000
	235	0.000	0.000	0.167	0.000	0.000	0.000	0.250	0.000
	239	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	247	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	255	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	277	0.000	0.000	0.000	0.000	0.167	0.000	0.000	0.000
	291	0.167	0.000	0.000	0.000	0.000	0.000	0.000	0.000


Allele Frequency for RHOS 29
1.200
1.000
0.800
0.600
0.400
0.200
0.000
Frequency
175
187
191
195
199
203
207
211
215
219
223
227
231
233
235
239
247
255
277
291
RHOS 29
Locus
INDIAN CK
MILL
PLUNGE
TWIN
CITY CK
CAJON
LYTLE
CATTLE CAN
EF SAN GAB
NF SAN GAB
FISH
WF SAN GAB
HAIN
PINE
MARVIN
BRIZZOLARI
STENNAR
SAN LUIS OBI
DAVY BROWN

APPENDIX B

PAIRWISE CHARTS

GEOGRAPHIC DISTANCE – 21 POPULATIONS

IND	MILL	PLNG	TWIN	CITY	CAJON	LYTLE	CATTL	EFSGR	NFSGR	FISH	WFSGR	HAIN	PINE	MARV	BRIZ	STEN	SLO	DB	CUY	MZNA	
0.000																					IND
40.359	0.000																				MILL
45.764	5.856	0.000																			PLUNGE
57.900	18.791	12.936	0.000																		TWIN
70.672	30.522	24.921	13.593	0.000																	CITY
72.453	34.736	28.923	16.113	12.955	0.000																CAJON
75.144	38.288	32.540	19.946	16.946	4.285	0.000															LYTLE
97.545	64.765	59.332	47.500	43.994	32.329	28.069	0.000														CATTL
125.228	94.534	89.204	77.512	73.376	62.239	58.011	30.016	0.000													EFSGR
125.228	94.534	89.204	77.512	73.376	62.239	58.011	30.016	0.000	0.000												NFSGR
106.239	76.757	71.659	60.658	58.604	46.407	42.121	15.115	19.083	19.083	0.000											FISH
125.228	94.535	89.204	77.512	73.376	62.239	58.011	30.016	0.000	0.000	19.083	0.000										WFSGR
139.280	110.240	104.040	93.628	89.895	78.586	74.330	46.265	16.569	16.569	35.595	16.569	0.000									HAIN
399.185	359.754	355.104	345.213	331.621	335.431	334.979	330.086	324.407	324.407	353.319	324.407	326.473	0.000								PINE
434.630	395.162	390.490	380.535	366.942	370.615	370.106	364.714	358.252	358.252	369.605	358.252	359.768	35.471	0.000							MAR
385.778	352.759	346.859	334.159	325.834	318.047	314.321	288.704	260.587	260.587	279.636	260.587	247.312	297.891	313.113	0.000						BRIZ
387.609	354.354	348.631	335.924	327.566	319.811	316.091	290.510	262.427	262.427	281.479	262.427	249.127	297.720	312.720	2.012	0.000					STEN
382.955	349.801	344.084	331.391	323.094	315.279	311.547	285.901	257.759	257.759	276.805	257.759	244.420	297.536	313.036	2.892	4.880	0.000				SLO
302.756	272.043	266.511	254.212	247.508	238.258	234.242	207.378	178.008	178.008	196.726	178.008	163.476	306.594	330.454	88.190	90.191	85.309	0.000			DB
303.143	270.017	264.315	251.656	243.580	235.549	231.785	206.052	177.985	177.985	197.048	177.985	164.901	279.089	302.412	82.656	84.473	79.849	28.053	0.000		CUYA
302.080	271.364	265.832	253.534	246.731	237.579	233.599	206.648	177.330	177.330	196.048	177.330	162.801	306.797	330.352	88.792	90.792	85.909	0.679	27.977	0.000	MZNA

PAIRWISE R_{ST} VALUE BELOW DIAGNOL AND p-VALUE ABOVE DIAGANOL FOR 21 POPULATIONS

IND	MILL	PLN	TWN	CITY	CAJON	LYTLE	CATTL	EFSG	NFSG	FISH	WFSG	HAIN	PINE	MARV	BRIZ	STEN	SLO	DB	CUY	MZNA	
0.000	0.002	0.002	0.001	0.005	0.001	0.004	0.004	0.001	0.019	0.001	0.003	0.002	0.001	0.001	0.001	0.001	0.001	0.003	0.001	0.001	**IND**
0.222	0.000	0.397	0.002	0.002	0.007	0.011	0.175	0.003	0.397	0.076	0.084	0.275	0.001	0.004	0.001	0.001	0.001	0.002	0.001	0.001	**MILL**
0.146	0.000	0.000	0.002	0.001	0.001	0.001	0.320	0.002	0.417	0.046	0.270	0.298	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	**PLNG**
0.335	0.208	0.098	0.000	0.001	0.001	0.001	0.012	0.001	0.319	0.001	0.074	0.004	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.001	**TWIN**
0.149	0.222	0.130	0.215	0.000	0.002	0.003	0.001	0.001	0.218	0.001	0.053	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	**CITY**
0.199	0.154	0.080	0.198	0.148	0.000	0.102	0.015	0.001	0.144	0.001	0.016	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	**CAJON**
0.158	0.158	0.112	0.232	0.124	0.018	0.000	0.078	0.001	0.161	0.002	0.053	0.017	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	**LYTLE**
0.315	0.067	0.015	0.205	0.250	0.161	0.092	0.000	0.034	0.466	0.358	0.457	0.460	0.001	0.013	0.002	0.001	0.002	0.001	0.012	0.001	**CATTL**
0.598	0.483	0.380	0.425	0.540	0.548	0.467	0.305	0.000	0.137	0.001	0.269	0.068	0.007	0.083	0.006	0.003	0.004	0.026	0.016	0.004	**EFSGR**
0.273	0.000	0.000	0.027	0.048	0.088	0.069	0.000	0.211	0.000	0.149	0.346	0.290	0.180	0.096	0.006	0.006	0.004	0.032	0.027	0.003	**NFSGR**
0.422	0.126	0.079	0.313	0.416	0.235	0.183	0.000	0.441	0.126	0.000	0.057	0.375	0.001	0.002	0.002	0.001	0.001	0.003	0.001	0.001	**FISH**
0.386	0.174	0.031	0.143	0.179	0.222	0.149	0.000	0.125	0.000	0.193	0.000	0.265	0.080	0.114	0.008	0.001	0.012	0.015	0.035	0.001	**WFSGR**
0.382	0.052	0.027	0.308	0.396	0.308	0.237	0.000	0.291	0.035	0.000	0.108	0.000	0.181	0.298	0.005	0.008	0.006	0.022	0.039	0.004	**HAIN**
0.415	0.292	0.278	0.397	0.413	0.472	0.430	0.279	0.287	0.088	0.382	0.131	0.097	0.000	0.456	0.006	0.024	0.041	0.040	0.082	0.002	**PINE**
0.414	0.319	0.349	0.418	0.521	0.549	0.518	0.283	0.193	0.195	0.362	0.190	0.051	0.000	0.000	0.003	0.004	0.001	0.016	0.024	0.001	**MARV**
0.727	0.728	0.637	0.771	0.744	0.712	0.654	0.746	0.772	0.650	0.770	0.720	0.679	0.267	0.439	0.000	0.101	0.109	0.102	0.263	0.058	**BRIZ**
0.736	0.766	0.627	0.768	0.759	0.714	0.653	0.787	0.805	0.718	0.794	0.786	0.731	0.309	0.354	0.148	0.000	0.457	0.293	0.441	0.150	**STENNER**
0.719	0.718	0.618	0.737	0.740	0.709	0.651	0.716	0.732	0.623	0.743	0.690	0.633	0.197	0.310	0.150	0.010	0.000	0.169	0.487	0.021	**SLO**
0.705	0.753	0.640	0.768	0.742	0.722	0.653	0.750	0.739	0.657	0.783	0.723	0.682	0.273	0.302	0.212	0.000	0.147	0.000	0.410	0.274	**DB**
0.683	0.727	0.596	0.746	0.721	0.683	0.611	0.738	0.754	0.637	0.763	0.726	0.665	0.226	0.299	0.070	0.000	0.000	0.000	0.000	0.412	**CUY**
0.762	0.809	0.664	0.807	0.773	0.730	0.657	0.834	0.860	0.785	0.835	0.842	0.808	0.454	0.531	0.170	0.112	0.276	0.067	0.000	0.000	**MZNA**

PAIRWISE F_{ST} VALUE BELOW DIAGNOL AND p-VALUE ABOVE DIAGANOL FOR 21 POPULATIONS

IND	MILL	PLNG	TWIN	CITY	CAJON	LYTLE	CATTL	EFSGR	NFSGR	FISH	WFSGR	HAIN	PINE	MARV	BRIZ	STEN	SLO	B	CUY	MZNA	
0.000	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	**IND**
0.304	0.000	0.001	0.001	0.001	0.001	0.001	0.003	0.002	0.001	0.001	0.001	0.002	0.001	0.001	0.001	0.001	0.002	0.002	0.001	0.001	**MILL**
0.202	0.088	0.000	0.001	0.001	0.001	0.001	0.004	0.005	0.003	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	**PLNG**
0.296	0.173	0.128	0.000	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	**TWIN**
0.237	0.165	0.081	0.160	0.000	0.001	0.001	0.004	0.025	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	**CITY**
0.208	0.147	0.073	0.138	0.066	0.000	0.001	0.001	0.001	0.063	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	**CAJON**
0.346	0.266	0.189	0.252	0.175	0.090	0.000	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	**LYTLE**
0.254	0.097	0.059	0.136	0.078	0.069	0.219	0.000	0.468	0.067	0.002	0.007	0.006	0.007	0.002	0.003	0.001	0.003	0.003	0.004	0.001	**CATTL**
0.286	0.133	0.080	0.178	0.072	0.093	0.254	0.000	0.000	0.061	0.001	0.056	0.020	0.018	0.010	0.006	0.003	0.005	0.037	0.019	0.005	**EFSGR**
0.312	0.177	0.089	0.144	0.102	0.044	0.217	0.061	0.094	0.000	0.002	0.025	0.032	0.008	0.008	0.006	0.003	0.005	0.024	0.031	0.001	**NFSGR**
0.319	0.237	0.127	0.160	0.192	0.138	0.301	0.097	0.193	0.134	0.000	0.005	0.005	0.001	0.001	0.001	0.001	0.001	0.002	0.001	0.001	**FISH**
0.396	0.294	0.208	0.286	0.224	0.240	0.393	0.186	0.170	0.309	0.328	0.000	0.025	0.007	0.006	0.004	0.006	0.005	0.017	0.033	0.005	**WFSGR**
0.291	0.248	0.157	0.242	0.185	0.140	0.309	0.139	0.187	0.195	0.188	0.396	0.000	0.004	0.002	0.004	0.008	0.003	0.017	0.034	0.003	**HAIN**
0.306	0.194	0.094	0.162	0.118	0.109	0.275	0.105	0.122	0.122	0.170	0.314	0.223	0.000	0.003	0.001	0.005	0.002	0.001	0.004	0.001	**PINE**
0.327	0.240	0.166	0.231	0.168	0.165	0.318	0.137	0.182	0.151	0.251	0.365	0.263	0.095	0.000	0.001	0.004	0.005	0.002	0.009	0.001	**MARV**
0.430	0.344	0.235	0.314	0.252	0.244	0.384	0.251	0.320	0.311	0.336	0.472	0.388	0.308	0.322	0.000	0.027	0.001	0.005	0.013	0.001	**BRIZ**
0.413	0.333	0.209	0.296	0.235	0.216	0.349	0.247	0.284	0.275	0.324	0.462	0.365	0.280	0.302	0.171	0.000	0.301	0.018	0.016	0.001	**STEN**
0.432	0.361	0.239	0.324	0.258	0.244	0.384	0.278	0.316	0.320	0.365	0.501	0.417	0.309	0.332	0.231	0.024	0.000	0.003	0.007	0.001	**SLO**
0.427	0.333	0.218	0.303	0.266	0.228	0.370	0.223	0.265	0.298	0.331	0.449	0.382	0.275	0.331	0.281	0.151	0.186	0.000	0.376	0.003	**DB**
0.371	0.315	0.217	0.293	0.242	0.219	0.360	0.217	0.266	0.252	0.322	0.457	0.333	0.265	0.309	0.284	0.135	0.199	0.036	0.000	0.048	**CUY**
0.425	0.414	0.295	0.360	0.334	0.296	0.419	0.345	0.424	0.421	0.416	0.567	0.450	0.404	0.431	0.330	0.238	0.248	0.205	0.101	0.000	**MZNA**

PAIRWISE R_{ST} VALUE BELOW DIAGNOL AND p-VALUE ABOVE DIAGANOL FOR SOUTHERN CALIFORNIA WATERSHEDS

SANTA ANA	SAN GABRIEL	
0.000	0.001	SANTA ANA
0.147	0.000	SAN GABRIEL

Rst Values below diagonal. Probability, P(rand >= data) based on 999 permutations is shown above diagonal.

PAIRWISE F_{ST} VALUE BELOW DIAGNOL AND p-VALUE ABOVE DIAGANOL FOR SOUTHERN CALIFORNIA WATERSHEDS

SANTA ANA	SAN GABRIEL	
0.000	0.001	SANTA ANA
0.053	0.000	SAN GABRIEL

Fst Values below diagonal. Probability, P(rand >= data) based on 999 permutations is shown above diagonal.

PAIRWISE R_{ST} VALUE BELOW DIAGNOL AND p-VALUE ABOVE DIAGANOL FOR SOUTHERN CALIFORNIA MOUNTAIN RANGES

SAN JACINTO	SAN BERNARDINO	SAN GABRIEL	
0.000	0.002	0.001	SAN JACINTO
0.135	0.000	0.001	SAN BERNARDINO
0.141	0.060	0.000	SAN GABRIEL

Rst Values below diagonal. Probability, P(rand >= data) based on 999 permutations is shown above diagonal.

PAIRWISE R_{ST} VALUE BELOW DIAGNOL AND p-VALUE ABOVE DIAGANOL FOR SOUTHERN CALIFORNIA MOUNTAIN RANGES

SAN JACINTO	SAN BERNARDINO	SAN GABRIEL	
0.000	0.001	0.001	SAN JACINTO
0.183	0.000	0.001	SAN BERNARDINO
0.215	0.059	0.000	SAN GABRIEL

Fst Values below diagonal. Probability, P(rand >= data) based on 999 permutations is shown above diagonal.

APPENDIX C

INPUT FILES

GENALEX MICROSATELLITE DATA FILE – 21 POPULATIONS

7	147	21	7	5	21	12	14	21	28	3	2	2	5	2	2	3	3	3	3	3	2	2	4	1	147
			INDIAN CK	MILL	PLUNGE	TWIN	CITY CK	CAJON	LYTLE	CATTLE CAN	EF SAN GAB	NF SAN GAB	FISH	WF SAN GAB	HAIN	PINE	MARVIN	BRIZZOLARI	STENNER	SAN LUIS OBI	DAVY BROWN	CUYAMA	MANZANA		CALIFO
Sample	Pop	RHOS 5		RHOS 9		RHOS 25		RHOS 23		RHOS 14		RHOS 8		RHOS 29			X	Y							
INDCK-01	INDIAN CK	232	232	137	137	144	180	236	236	270	274	204	212	203	219		33.761964	-116.883911							
INDCK-02	INDIAN CK	232	236	137	157	148	148	236	236	270	278	192	212	199	199		33.761964	-116.883911							
INDCK-03	INDIAN CK	232	256	149	157	148	148	236	236	250	270	204	204	219	235		33.761964	-116.883911							
INDCK-04	INDIAN CK	232	232	145	149	148	148	236	236	270	274	192	204	235	235		33.761964	-116.883911							
INDCK-05	INDIAN CK	232	232	137	137	128	148	236	244	262	266	192	204	199	219		33.761964	-116.883911							
INDCK-06	INDIAN CK	232	232	149	157	148	148	236	236	270	270	192	192	195	199		33.761964	-116.883911							
INDCK-07	INDIAN CK	256	256	137	137	184	216	236	236	270	270	192	192	199	219		33.761964	-116.883911							
MILL-1	MILL	252	256	137	141	156	160	232	252	282	282	192	200	219	227		34.077789	-117.09945							
MILL-2	MILL	252	264	137	141	160	160	232	252	274	282	216	216	203	207		34.077789	-117.09945							
MILL-3	MILL	264	276	137	137	160	160	228	252	282	282	192	264	207	215		34.077789	-117.09945							
MILL-4	MILL	236	276	137	137	140	172	232	232	274	282	200	220	199	215		34.077789	-117.09945							
MILL-5	MILL	228	272	133	137	140	160	252	252	274	282	200	216	203	215		34.077789	-117.09945							
PC-01	PLUNGE	252	252	133	137	140	188	228	228	238	238	192	196	223	227		34.108917	-117.15075							

PC-02	PLUNGE	236	280	137	137	116	116	256	256	238	238	192	196	223	227	34.108917	-117.15075
PC-03	PLUNGE	228	264	137	141	180	192	248	248	238	238	204	220	227	227	34.108917	-117.15075
PC-04	PLUNGE	252	264	137	157	140	152	224	248	302	302	200	204	203	215	34.108917	-117.15075
PC-05	PLUNGE	252	276	141	157	140	160	256	256	278	286	204	224	223	223	34.108917	-117.15075
PC-06	PLUNGE	252	280	137	157	140	180	256	256	254	282	204	268	195	219	34.108917	-117.15075
PC-07	PLUNGE	264	280	137	137	140	180	232	232	274	258	220	268	203	227	34.108917	-117.15075
PC-08	PLUNGE	252	280	133	145	152	152	228	256	286	286	192	204	231	203	34.108917	-117.15075
PC-09	PLUNGE	252	280	137	137	152	192	232	232	254	274	196	204	203	227	34.108917	-117.15075
PC-10	PLUNGE	276	276	133	137	140	140	228	228	254	266	192	200	203	247	34.108917	-117.15075
PC-11	PLUNGE	228	268	137	141	168	168	244	244	278	282	200	204	227	277	34.108917	-117.15075
PC-12	PLUNGE	228	268	133	137	168	192	248	248	266	278	192	220	195	203	34.108917	-117.15075
PC-13	PLUNGE	236	280	137	145	136	184	248	248	246	274	200	200	215	227	34.108917	-117.15075
PC-14	PLUNGE	252	280	137	137	168	180	252	256	274	274	196	204	215	215	34.108917	-117.15075
PC-15	PLUNGE	252	276	141	141	168	168	228	256	290	290	200	200	175	175	34.108917	-117.15075
PC-16	PLUNGE	236	252	141	161	148	168	232	252	282	282	192	200	199	203	34.108917	-117.15075
PLNG-1	PLUNGE	236	276	137	141	140	184	248	248	266	290	200	204	203	203	34.108917	-117.15075
PLNG-2	PLUNGE	252	268	157	157	140	172	228	256	258	290	204	226	215	219	34.1089	-117.1

												8			7	5075
PLNG-3	PLUNGE	276	276	137	145	140	184	244	244	266	286	192	204	219	227	34.1089177	-117.15075
PLNG-4	PLUNGE	264	276	137	161	172	172	244	244	266	266	204	220	215	219	34.1089177	-117.15075
PLUN-1	PLUNGE	276	276	157	161	152	172	256	256	266	266	196	224	203	219	34.1089177	-117.15075
TWIN-01	TWIN	256	272	133	133	152	152	264	264	262	262	184	196	219	227	34.174722	-117.266667
TWIN-02	TWIN	224	224	133	137	192	192	232	264	262	290	184	200	211	223	34.174722	-117.266667
TWIN-03	TWIN	224	284	153	153	192	196	264	268	262	262	200	212	211	211	34.174722	-117.266667
TWIN-04	TWIN	224	264	133	137	160	160	256	268	266	270	200	212	211	223	34.174722	-117.266667
TWIN-05	TWIN	264	264	133	133	156	168	228	264	290	290	184	200	211	219	34.174722	-117.266667
TWIN-06	TWIN	260	292	133	133	156	168	256	268	266	290	200	200	211	227	34.174722	-117.266667
TWIN-07	TWIN	260	284	121	133	160	160	260	260	274	282	200	200	219	219	34.174722	-117.266667
TWIN-08	TWIN	252	252	133	133	192	192	252	264	262	282	200	200	211	211	34.174722	-117.266667
TWIN-09	TWIN	256	268	137	141	156	192	264	264	254	274	196	200	211	227	34.174722	-117.266667
TWIN-10	TWIN	224	264	137	145	156	168	264	264	262	274	200	200	211	211	34.174722	-117.266667
TWIN-11	TWIN	256	292	133	133	156	192	252	252	278	278	200	200	203	215	34.174722	-117.266667
TWIN-12	TWIN	252	280	133	137	156	192	236	264	270	270	200	200	211	227	34.174722	-117.266667
CC-01	CITYCK	236	244	149	161	152	160	228	228	270	270	184	184	235	235	34.292526	-117.306135

CC-02	CITY CK	244	252	137	145	140	164	260	260	278	282	184	200	227	235	34.2925266	-117.306135
CC-03	CITY CK	236	236	137	149	140	156	244	256	274	274	188	196	227	227	34.2925266	-117.306135
CC-04	CITY CK	244	276	145	157	156	172	244	244	274	278	196	196	231	235	34.2925266	-117.306135
CC-05	CITY CK	244	276	141	165	164	172	256	256	270	274	184	188	219	227	34.2925266	-117.306135
CC-06	CITY CK	232	272	129	149	164	168	244	260	274	294	208	220	227	227	34.2925266	-117.306135
CC-07	CITY CK	244	244	141	145	152	152	244	244	274	274	184	188	227	231	34.2925266	-117.306135
CC-08	CITY CK	232	272	141	149	168	168	244	244	258	274	196	204	227	227	34.2925266	-117.306135
CC-09	CITY CK	240	276	137	153	168	168	244	244	274	278	200	208	227	227	34.2925266	-117.306135
CC-10	CITY CK	240	276	141	153	156	200	244	256	274	274	204	220	227	227	34.2925266	-117.306135
CC-11	CITY CK	244	244	125	133	156	168	240	248	278	278	188	220	227	227	34.2925266	-117.306135
CCWF-01	CITY CK	236	244	149	161	164	208	228	228	274	274	188	192	227	235	34.2925266	-117.306135
FORK-1	CITY CK	224	224	129	137	164	164	236	256	282	294	188	196	235	247	34.2925266	-117.306135
FORK-2	CITY CK	224	244	137	145	132	172	208	208	254	294	200	212	199	227	34.2925266	-117.306135
CAJO-01	CAJON	264	276	141	141	172	172	232	252	254	270	196	264	199	231	34.2328111	-117.427183
CAJO-02	CAJON	252	252	137	137	156	156	208	224	266	282	196	236	203	231	34.2328111	-117.427183
CAJO-03	CAJON	268	280	125	157	168	172	228	228	254	270	196	196	215	247	34.2328111	-117.427183
CAJO-04	CAJON	264	268	133	153	216	220	240	256	254	274	196	211	191	227	34.2328	-117.4

												2			1	27183	
CAJO-05	CAJON	276	276	125	157	168	172	228	228	254	270	196	196	215	247	34.23281	-117.427183
CAJO-06	CAJON	224	264	133	145	156	172	232	232	254	270	212	212	227	247	34.23281	-117.427183
CAJO-07	CAJON	220	252	145	149	176	176	208	228	254	270	204	208	211	235	34.23281	-117.427183
CAJO-08	CAJON	224	224	141	145	132	172	224	240	250	254	188	196	223	231	34.23281	-117.427183
CAJO-09	CAJON	280	280	141	141	132	216	228	232	254	274	208	208	199	227	34.23281	-117.427183
CAJO-10	CAJON	272	276	133	141	176	176	232	244	254	266	212	216	191	227	34.23281	-117.427183
CAJO-11	CAJON	260	264	145	145	164	164	228	240	274	294	196	204	199	227	34.23281	-117.427183
CAJO-12	CAJON	264	280	133	145	132	176	228	244	266	274	196	212	227	239	34.23281	-117.427183
CAJO-13	CAJON	252	264	141	145	168	216	212	228	250	270	212	216	215	227	34.23281	-117.427183
CAJO-14	CAJON	224	248	133	145	164	180	228	232	270	250	204	212	199	211	34.23281	-117.427183
CAJO-15	CAJON	252	272	137	141	176	176	236	244	270	274	192	196	227	239	34.23281	-117.427183
CAJO-16	CAJON	248	268	129	141	132	168	244	244	266	270	188	216	247	247	34.23281	-117.427183
CAJO-17	CAJON	248	248	141	141	168	168	228	236	274	278	188	188	199	227	34.23281	-117.427183
CAJW-1	CAJON	224	264	133	141	176	176	232	232	254	266	196	212	199	227	34.23281	-117.427183
CAJW-2	CAJON	220	272	133	145	172	212	244	244	254	270	188	196	227	235	34.23281	-117.427183
CAJW-3	CAJON	248	276	133	141	176	180	232	236	266	278	196	212	191	227	34.23281	-117.427183

CAJW-4	CAJON	248	276	133	141	184	184	208	236	270	270	196	212	199	227	34.23281	-117.427183
LCKC-01	LYTLE CK	220	276	141	141	156	156	236	236	254	270	188	196	227	247	34.28583	-117.473517
LCKC-02	LYTLE CK	252	276	141	141	156	176	232	232	270	274	264	264	247	247	34.28583	-117.473517
LCKC-03	LYTLE CK	220	264	141	141	156	216	228	228	254	254	212	212	211	247	34.28583	-117.473517
LCKC-04	LYTLE CK	264	276	141	141	156	156	228	236	254	254	188	196	211	211	34.28583	-117.473517
LCKC-05	LYTLE CK	264	268	141	141	176	176	228	228	254	270	188	188	211	227	34.28583	-117.473517
LCKC-06	LYTLE CK	264	276	141	141	156	156	228	228	254	274	208	260	227	247	34.28583	-117.473517
LCKC-07	LYTLE CK	220	276	141	141	148	152	228	232	254	274	188	188	227	227	34.28583	-117.473517
LCKC-08	LYTLE CK	220	276	145	145	212	212	236	236	254	254	188	188	227	227	34.28583	-117.473517
LCKC-09	LYTLE CK	220	276	141	141	156	216	228	236	254	274	212	264	227	227	34.28583	-117.473517
LCKC-10	LYTLE CK	220	276	141	141	156	216	228	228	254	254	196	212	227	247	34.28583	-117.473517
LCK-11	LYTLE CK	220	276	141	141	176	216	228	232	254	254	188	188	211	227	34.28583	-117.473517
LCK-12	LYTLE CK	220	252	141	141	156	156	228	232	254	254	188	188	227	227	34.28583	-117.473517
LCK-13	LYTLE CK	220	268	137	141	156	216	228	236	270	274	188	196	247	247	34.28583	-117.473517
LCK-14	LYTLE CK	220	276	137	141	156	176	228	232	254	254	196	208	211	227	34.28583	-117.473517
LCK-15	LYTLE CK	252	260	141	145	172	176	228	236	254	254	196	256	211	227	34.28583	-117.473517
LCK-16	LYTLE CK	220	276	137	141	156	216	228	228	254	270	196	119	211	211	34.2858	-117.4

												6			3	73517
LCK-17	LYTLE CK	220	276	137	141	156	216	232	232	254	270	264	264	211	227	34.2285833	-117.473517
LCK-18	LYTLE CK	252	264	141	141	156	156	228	228	270	270	212	264	227	227	34.2285833	-117.473517
LCK-19	LYTLE CK	252	260	145	145	156	156	228	236	254	270	188	188	227	277	34.2285833	-117.473517
LCK-20	LYTLE CK	220	276	141	141	156	176	232	236	254	254	196	212	227	247	34.2285833	-117.473517
LCK-21	LYTLE CK	260	276	141	141	156	156	228	236	254	254	264	264	211	227	34.2285833	-117.473517
LCK-22	LYTLE CK	264	276	137	141	156	176	220	232	254	274	196	196	211	227	34.2285833	-117.473517
LCK-23	LYTLE CK	220	264	141	141	156	156	228	232	254	274	188	212	227	247	34.2285833	-117.473517
LCK-24	LYTLE CK	220	268	141	141	156	156	228	232	274	274	196	196	211	227	34.2285833	-117.473517
LCK-25	LYTLE CK	220	220	133	141	156	156	232	232	254	282	188	212	247	247	34.2285833	-117.473517
LCK-26	LYTLE CK	264	276	141	141	156	176	232	236	254	274	188	188	227	227	34.2285833	-117.473517
LCK-27	LYTLE CK	264	276	141	141	152	152	224	228	254	270	196	212	227	247	34.2285833	-117.473517
LCK-28	LYTLE CK	220	280	141	141	172	172	220	228	254	274	196	260	211	247	34.2285833	-117.473517
CATCR-01	CATTLE CAN	244	284	125	137	160	164	244	248	258	262	212	216	219	231	34.23	-117.778832
CATCR-02	CATTLE CAN	228	252	125	161	152	156	252	268	278	278	228	228	203	207	34.23	-117.778832
CATCR-03	CATTLE CAN	252	264	125	141	160	172	232	256	270	274	224	228	231	239	34.23	-117.778832
EFSGR-01	EF SAN GAB	240	244	137	141	160	164	276	260	266	298	220	228	227	239	34.258056	-118.103611

EFSGR-02	EF SAN GAB	244	252	137	145	160	164	276	340	278	298	224	232	199	231	34.25805	-118.103611
NFSGR-01	NF SAN GAB	264	272	133	141	152	172	232	256	266	294	196	208	187	235	34.25805	-118.103611
NFSGR-02	NF SAN GAB	240	264	129	145	156	172	252	264	266	294	196	232	235	255	34.25805	-118.103611
FISH-1	FISH	252	260	133	133	152	168	248	252	266	270	212	228	203	231	34.16937	-117.925931
FISH-2	FISH	276	288	133	133	152	172	268	268	262	274	224	244	223	231	34.16937	-117.925931
FISH-3	FISH	220	264	133	137	152	152	220	220	270	270	224	244	203	203	34.16937	-117.925931
FISH-4	FISH	260	284	133	133	152	176	248	248	266	266	208	212	203	203	34.16937	-117.925931
FISH-5	FISH	260	260	129	133	152	172	248	248	262	262	240	244	215	219	34.16937	-117.925931
WFSGR-01	WF SAN GAB	252	260	137	153	160	164	260	260	278	278	192	240	223	235	34.25805	-118.103611
WFSGR-01	WF SAN GAB	252	260	137	153	160	164	260	260	278	278	192	240	223	235	34.25805	-118.103611
HAIN-1	HAIN	260	260	117	145	148	148	232	276	270	274	216	236	203	203	34.236788	-118.282027
HAIN-2	HAIN	260	268	145	145	132	152	232	276	270	274	204	236	219	219	34.236788	-118.282027
PINE-1	PINE	280	280	129	137	128	132	244	276	314	314	180	172	227	227	37.172778	-118.259444
PINE-2	PINE	272	280	193	197	128	128	248	280	238	238	176	184	175	291	37.172778	-118.259444
PINE-3	PINE	272	276	133	133	128	132	244	280	262	294	200	228	215	215	37.172778	-118.259444
MARV-1	MARVIN	248	256	121	169	128	128	264	296	258	258	172	176	207	191	37.470783	-118.402549
MARV-2	MARVIN	268	272	129	129	128	128	256	300	258	274	168	200	187	187	37.4707	-118.4

												8			3	02549
MARV-3	MARVIN	244	284	129	225	128	128	256	304	274	282	172	298	199	199	37.47078 3	-118.402549
BRIZ-1	BRIZZOLARI	280	292	177	177	108	108	196	248	274	282	160	160	231	231	35.33613 68	-120.651393
BRIZ-2	BRIZZOLARI	292	308	177	177	112	112	256	256	230	242	160	160	223	223	35.33613 68	-120.651393
BRIZ-3	BRIZZOLARI	248	292	177	177	108	108	196	196	274	274	160	160	235	191	35.33613 68	-120.651393
STEN-1	STENNER	260	284	173	181	108	108	256	256	242	242	160	160	195	227	35.32760 3	-120.665568
STEN-2	STENNER	264	276	177	177	116	116	196	220	242	242	160	160	191	227	35.32760 3	-120.665568
STEN-3	STENNER	276	276	145	177	108	112	256	264	242	242	160	160	199	199	35.32760 3	-120.665568
SLO-1	SAN LUIS OBI	272	272	165	165	108	108	256	256	242	242	160	160	199	199	35.29758 5	-120.626288
SLO-2	SAN LUIS OBI	292	300	137	145	108	108	256	300	242	242	160	160	191	227	35.29758 5	-120.626288
SLO-3	SAN LUIS OBI	284	288	161	177	108	108	216	224	242	242	160	160	227	277	35.29758 5	-120.626288
DBC-01	DAVY BROWN	252	256	189	189	108	112	276	280	242	242	160	160	175	223	34.77207 8	-119.943624
DBC-02	DAVY BROWN	252	252	173	185	108	112	204	228	242	242	160	160	233	233	34.77207 8	-119.943624
CUYR-01	CUYAMA	256	276	149	149	112	112	200	236	242	242	160	160	219	219	35.00464 5	-119.82442
CUYR-02	CUYAMA	252	272	185	185	108	112	232	276	242	242	160	160	187	235	35.00464 5	-119.82442
MZNA-01	MANZANA	256	292	177	185	108	108	196	196	242	242	160	160	219	219	34.77043 6	-119.936469
MZNA-02	MANZANA	248	252	149	173	108	112	232	236	242	242	160	160	219	219	34.77043 6	-119.936469

MNCR-01	MANZANA	256	256	173	189	108	108	232	236	242	242	160	160	219	219	34.77043636	-119.936469
MNCR-02	MANZANA	248	292	173	189	108	112	236	236	242	242	160	160	219	219	34.77043636	-119.936469

GENALEX MICROSATELLITE DATA FILE – SOUTHERN CALIFRONIA WATERSHEDS

7	123	2	108	15	1	123													
			SANTA ANA	SAN GABRIEL		SOCAL													
Sample	Pop	RHOS 5		RHOS 9		RHOS 25		RHOS 23		RHOS 14		RHOS 8		RHOS 29		X	Y
INDCK-01	SANTA ANA	232	232	137	137	144	180	236	236	270	274	204	212	203	219	33.761964	-116.883911
INDCK-02	SANTA ANA	232	236	137	157	148	148	236	236	270	278	192	212	199	199	33.761964	-116.883911
INDCK-03	SANTA ANA	232	256	149	157	148	148	236	236	250	270	204	204	219	235	33.761964	-116.883911
INDCK-04	SANTA ANA	232	232	145	149	148	148	236	236	270	274	192	204	235	235	33.761964	-116.883911
INDCK-05	SANTA ANA	232	232	137	137	128	148	236	244	262	266	192	204	199	219	33.761964	-116.883911
INDCK-06	SANTA ANA	232	232	149	157	148	148	236	236	270	270	192	192	195	199	33.761964	-116.883911
INDCK-07	SANTA ANA	256	256	137	137	184	216	236	236	270	270	192	192	199	219	33.761964	-116.883911
MILL-1	SANTA ANA	252	256	137	141	156	160	232	252	282	282	192	200	219	227	34.077789	-117.09945
MILL-2	SANTA ANA	252	264	137	141	160	160	232	252	274	282	216	216	203	207	34.077789	-117.09945
MILL-3	SANTA ANA	264	276	137	137	160	160	228	252	282	282	192	264	207	215	34.077789	-117.09945
MILL-4	SANTA	236	276	137	13	140	17	232	23	274	28	200	22	199	21	34.0777	-

	ANA				7		2		2		2		0		5	89	117.09945
MILL-5	SANTA ANA	228	272	133	137	140	160	252	252	274	282	200	216	203	215	34.077789	-117.09945
PC-01	SANTA ANA	252	252	133	137	140	188	228	228	238	238	192	196	223	227	34.108917	-117.15075
PC-02	SANTA ANA	236	280	137	137	116	116	256	256	238	238	192	196	223	227	34.108917	-117.15075
PC-03	SANTA ANA	228	264	137	141	180	192	248	248	238	238	204	220	227	227	34.108917	-117.15075
PC-04	SANTA ANA	252	264	137	157	140	152	224	248	302	302	200	204	203	215	34.108917	-117.15075
PC-05	SANTA ANA	252	276	141	157	140	160	256	256	278	286	204	224	223	223	34.108917	-117.15075
PC-06	SANTA ANA	252	280	137	157	140	180	256	256	254	282	204	268	195	219	34.108917	-117.15075
PC-07	SANTA ANA	264	280	137	137	140	180	232	232	274	258	220	268	203	227	34.108917	-117.15075
PC-08	SANTA ANA	252	280	133	145	152	152	228	256	286	286	192	204	231	203	34.108917	-117.15075
PC-09	SANTA ANA	252	280	137	137	152	192	232	232	254	274	196	204	203	227	34.108917	-117.15075
PC-10	SANTA ANA	276	276	133	137	140	140	228	228	254	266	192	220	203	247	34.108917	-117.15075
PC-11	SANTA ANA	228	268	137	141	168	168	244	244	278	282	200	204	227	277	34.108917	-117.15075
PC-12	SANTA	228	268	133	13	168	19	248	24	266	27	192	22	195	20	34.1089	-

	ANA				7		2		8		8		0		3	17	117.15075
PC-13	SANTA ANA	236	280	137	145	136	184	248	248	246	274	200	200	215	227	34.108917	-117.15075
PC-14	SANTA ANA	252	280	137	137	168	180	252	256	274	274	196	204	215	215	34.108917	-117.15075
PC-15	SANTA ANA	252	276	141	141	168	168	228	256	290	290	200	200	175	175	34.108917	-117.15075
PC-16	SANTA ANA	236	252	141	161	148	168	232	252	282	282	192	200	199	203	34.108917	-117.15075
PLNG-1	SANTA ANA	236	276	137	141	140	184	248	248	266	290	200	204	203	203	34.108917	-117.15075
PLNG-2	SANTA ANA	252	268	157	157	140	172	228	256	258	290	204	268	215	219	34.108917	-117.15075
PLNG-3	SANTA ANA	276	276	137	145	140	184	244	244	266	286	192	204	219	227	34.108917	-117.15075
PLNG-4	SANTA ANA	264	276	137	161	172	172	244	244	266	266	204	220	215	219	34.108917	-117.15075
PLUN-1	SANTA ANA	276	276	157	161	152	172	256	256	266	266	196	224	203	219	34.108917	-117.15075
TWIN-01	SANTA ANA	256	272	133	133	152	152	264	264	262	262	184	196	219	227	34.174722	-117.266667
TWIN-02	SANTA ANA	224	224	133	137	192	192	232	264	262	290	184	200	211	223	34.174722	-117.266667
TWIN-03	SANTA ANA	224	284	153	153	192	196	264	268	262	262	200	212	211	211	34.174722	-117.266667
TWIN-	SANTA	224	264	133	13	160	16	256	26	266	27	200	21	211	22	34.1747	-

04	ANA				7		0		8		0		2		3	22	117.266667
TWIN-05	SANTA ANA	264	264	133	133	156	168	228	264	290	290	184	200	211	219	34.174722	-117.266667
TWIN-06	SANTA ANA	260	292	133	133	156	168	256	268	266	290	200	200	211	227	34.174722	-117.266667
TWIN-07	SANTA ANA	260	284	121	133	160	160	260	260	274	282	200	200	219	219	34.174722	-117.266667
TWIN-08	SANTA ANA	252	252	133	133	192	192	252	264	262	282	200	200	211	211	34.174722	-117.266667
TWIN-09	SANTA ANA	256	268	137	141	156	192	264	264	254	274	196	200	211	227	34.174722	-117.266667
TWIN-10	SANTA ANA	224	264	137	145	156	168	264	264	262	274	200	200	211	211	34.174722	-117.266667
TWIN-11	SANTA ANA	256	292	133	133	156	192	252	252	278	278	200	200	203	215	34.174722	-117.266667
TWIN-12	SANTA ANA	252	280	133	137	156	192	236	264	270	270	200	200	211	227	34.174722	-117.266667
CC-01	SANTA ANA	236	244	149	161	152	160	228	228	270	270	184	184	235	235	34.292526	-117.306135
CC-02	SANTA ANA	244	252	137	145	140	164	260	260	278	282	184	200	227	235	34.292526	-117.306135
CC-03	SANTA ANA	236	236	137	149	140	156	244	256	274	274	188	196	227	227	34.292526	-117.306135
CC-04	SANTA ANA	244	276	145	157	156	172	244	244	274	278	196	196	231	235	34.292526	-117.306135
CC-05	SANTA	244	276	141	16	164	17	256	25	270	27	184	18	219	22	34.2925	-

	ANA				5		2		6		4		8		7	26	117.306135
CC-06	SANTA ANA	232	272	129	149	164	168	244	260	274	294	208	220	227	227	34.292526	-117.306135
CC-07	SANTA ANA	244	244	141	145	152	152	244	244	274	274	184	188	227	231	34.292526	-117.306135
CC-08	SANTA ANA	232	272	141	149	168	168	244	244	258	274	196	204	227	227	34.292526	-117.306135
CC-09	SANTA ANA	240	276	137	153	168	168	244	244	274	278	200	208	227	227	34.292526	-117.306135
CC-10	SANTA ANA	240	276	141	153	156	200	244	256	274	274	204	220	227	227	34.292526	-117.306135
CC-11	SANTA ANA	244	244	125	133	156	168	240	248	278	278	188	220	227	227	34.292526	-117.306135
CCWF-01	SANTA ANA	236	244	149	161	164	208	228	228	274	274	188	192	227	235	34.292526	-117.306135
FORK-1	SANTA ANA	224	224	129	137	164	164	236	256	282	294	188	196	235	247	34.292526	-117.306135
FORK-2	SANTA ANA	224	244	137	145	132	172	208	208	254	294	200	212	199	227	34.292526	-117.306135
CAJO-01	SANTA ANA	264	276	141	141	172	172	232	252	254	270	196	264	199	231	34.232811	-117.427183
CAJO-02	SANTA ANA	252	252	137	137	156	156	208	224	266	282	196	236	203	231	34.232811	-117.427183
CAJO-03	SANTA ANA	268	280	125	157	168	172	228	228	254	270	196	196	215	247	34.232811	-117.427183
CAJO-04	SANTA	264	268	133	15	216	22	240	25	254	27	196	21	191	22	34.2328	-

	ANA				3		0		6		4		2		7	11	117.427183
CAJO-05	SANTA ANA	276	276	125	157	168	172	228	228	254	270	196	196	215	247	34.232811	-117.427183
CAJO-06	SANTA ANA	224	264	133	145	156	172	232	232	254	270	212	212	227	247	34.232811	-117.427183
CAJO-07	SANTA ANA	220	252	145	149	176	176	208	228	254	270	204	208	211	235	34.232811	-117.427183
CAJO-08	SANTA ANA	224	224	141	145	132	172	224	240	250	254	188	196	223	231	34.232811	-117.427183
CAJO-09	SANTA ANA	280	280	141	141	132	216	228	232	254	274	208	208	199	227	34.232811	-117.427183
CAJO-10	SANTA ANA	272	276	133	141	176	176	232	244	254	266	212	216	191	227	34.232811	-117.427183
CAJO-11	SANTA ANA	260	264	145	145	164	164	228	240	274	294	196	204	199	227	34.232811	-117.427183
CAJO-12	SANTA ANA	264	280	133	145	132	176	228	244	266	274	196	212	227	239	34.232811	-117.427183
CAJO-13	SANTA ANA	252	264	141	145	168	216	212	228	250	270	212	216	215	227	34.232811	-117.427183
CAJO-14	SANTA ANA	224	248	133	145	164	180	228	232	270	250	204	212	199	211	34.232811	-117.427183
CAJO-15	SANTA ANA	252	272	137	141	176	176	236	244	270	274	192	196	227	239	34.232811	-117.427183
CAJO-16	SANTA ANA	248	268	129	141	132	168	244	244	266	270	188	216	247	247	34.232811	-117.427183
CAJO-17	SANTA	248	248	141	14	168	16	228	23	274	27	188	18	199	22	34.2328	-

	ANA				1		8		6		8		8		7	11	117.427183
CAJW-1	SANTA ANA	224	264	133	141	176	176	232	232	254	266	196	212	199	227	34.232811	-117.427183
CAJW-2	SANTA ANA	220	272	133	145	172	212	244	244	254	270	188	196	227	235	34.232811	-117.427183
CAJW-3	SANTA ANA	248	276	133	141	176	180	232	236	266	278	196	212	191	227	34.232811	-117.427183
CAJW-4	SANTA ANA	248	276	133	141	184	184	208	236	270	270	196	212	199	227	34.232811	-117.427183
LCKC-01	SANTA ANA	220	276	141	141	156	156	236	236	254	270	188	196	227	247	34.228583	-117.473517
LCKC-02	SANTA ANA	252	276	141	141	156	176	232	232	270	274	264	264	247	247	34.228583	-117.473517
LCKC-03	SANTA ANA	220	264	141	141	156	216	228	228	254	254	212	212	211	247	34.228583	-117.473517
LCKC-04	SANTA ANA	264	276	141	141	156	156	228	236	254	254	188	196	211	211	34.228583	-117.473517
LCKC-05	SANTA ANA	264	268	141	141	176	176	228	228	254	270	188	188	211	227	34.228583	-117.473517
LCKC-06	SANTA ANA	264	276	141	141	156	156	228	228	254	274	208	260	227	247	34.228583	-117.473517
LCKC-07	SANTA ANA	220	276	141	141	148	152	228	232	254	274	188	188	227	227	34.228583	-117.473517
LCKC-08	SANTA ANA	220	276	145	145	212	212	236	236	254	254	188	188	227	227	34.228583	-117.473517
LCKC-09	SANTA	220	276	141	14	156	21	228	23	254	27	212	26	227	22	34.2285	-

	ANA				1		6		6		4		4		7	83	117.473517
LCKC-10	SANTA ANA	220	276	141	141	156	216	228	228	254	254	196	212	227	247	34.228583	-117.473517
LCK-11	SANTA ANA	220	276	141	141	176	216	228	232	254	254	188	188	211	227	34.228583	-117.473517
LCK-12	SANTA ANA	220	252	141	141	156	156	228	232	254	254	188	188	227	227	34.228583	-117.473517
LCK-13	SANTA ANA	220	268	137	141	156	216	228	236	270	274	188	196	247	247	34.228583	-117.473517
LCK-14	SANTA ANA	220	276	137	141	156	176	228	232	254	254	196	208	211	227	34.228583	-117.473517
LCK-15	SANTA ANA	252	260	141	145	172	176	228	236	254	254	196	256	211	227	34.228583	-117.473517
LCK-16	SANTA ANA	220	276	137	141	156	216	228	228	254	270	196	196	211	211	34.228583	-117.473517
LCK-17	SANTA ANA	220	276	137	141	156	216	232	232	254	270	264	264	211	227	34.228583	-117.473517
LCK-18	SANTA ANA	252	264	141	141	156	156	228	228	270	270	212	264	227	227	34.228583	-117.473517
LCK-19	SANTA ANA	252	260	145	145	156	156	228	236	254	270	188	188	227	277	34.228583	-117.473517
LCK-20	SANTA ANA	220	276	141	141	156	176	232	236	254	254	196	212	227	247	34.228583	-117.473517
LCK-21	SANTA ANA	260	276	141	141	156	156	228	236	254	254	264	264	211	227	34.228583	-117.473517
LCK-22	SANTA	264	276	137	14	156	17	220	23	254	27	196	19	211	22	34.2285	-

	ANA				1		6		2		4		6		7	83	117.473517
LCK-23	SANTA ANA	220	264	141	141	156	156	228	232	254	274	188	212	227	247	34.228583	-117.473517
LCK-24	SANTA ANA	220	268	141	141	156	156	228	232	274	274	196	196	211	227	34.228583	-117.473517
LCK-25	SANTA ANA	220	220	133	141	156	156	232	232	254	282	188	212	247	247	34.228583	-117.473517
LCK-26	SANTA ANA	264	276	141	141	156	176	232	236	254	274	188	188	227	227	34.228583	-117.473517
LCK-27	SANTA ANA	264	276	141	141	152	152	224	228	254	270	196	212	227	247	34.228583	-117.473517
LCK-28	SANTA ANA	220	280	141	141	172	172	220	228	254	274	196	260	211	247	34.228583	-117.473517
CATCR-01	SAN GABRIEL	244	284	125	137	160	164	244	248	258	262	212	216	219	231	34.23	-117.778832
CATCR-02	SAN GABRIEL	228	252	125	161	152	156	252	268	278	278	228	228	203	207	34.23	-117.778832
CATCR-03	SAN GABRIEL	252	264	125	141	160	172	232	256	270	274	224	228	231	239	34.23	-117.778832
EFSGR-01	SAN GABRIEL	240	244	137	141	160	164	276	260	266	298	220	228	227	239	34.258056	-118.103611
EFSGR-02	SAN GABRIEL	244	252	137	145	160	164	276	340	278	298	224	232	199	231	34.258056	-118.103611
NFSGR-01	SAN GABRIEL	264	272	133	141	152	172	232	256	266	294	196	208	187	235	34.258056	-118.103611
NFSGR-	SAN	240	264	129	14	156	17	252	26	266	29	196	23	235	25	34.2580	-

02	GABRIEL				5		2		4		4		2		5	56	118.103611
FISH-1	SAN GABRIEL	252	260	133	133	152	168	248	252	266	270	212	228	203	231	34.169372	-117.925931
FISH-2	SAN GABRIEL	276	288	133	133	152	172	268	268	262	274	224	244	223	231	34.169372	-117.925931
FISH-3	SAN GABRIEL	220	264	133	137	152	152	220	220	270	270	224	244	203	203	34.169372	-117.925931
FISH-4	SAN GABRIEL	260	284	133	133	152	176	248	248	266	266	208	212	203	203	34.169372	-117.925931
FISH-5	SAN GABRIEL	260	260	129	133	152	172	248	248	262	262	240	244	215	219	34.169372	-117.925931
WFSGR-01	SAN GABRIEL	252	260	137	153	160	164	260	260	278	278	192	240	223	235	34.258056	-118.103611
HAIN-1	SAN GABRIEL	260	260	117	145	148	148	232	276	270	274	216	236	203	203	34.236788	-118.282027
HAIN-2	SAN GABRIEL	260	268	145	145	132	152	232	276	270	274	204	236	219	219	34.236788	-118.282027

GENALEX MICROSATELLITE DATA FILE – MOUNTAIN RANGES

0.2	123	3	7	52	64	1	123											
			SAN JACINTO	SAN BERNARDINO	SAN GABRIEL													
Sample	Pop	Rhos 5		Rhos 9		Rhos 25		Rhos 23		Rhos 14		Rhos 8		Rhos 29		X	Y	
INDCK-01	SAN JACINTO	232	232	137	137	144	180	236	236	270	274	204	212	203	219	33.761964	-116.883911	
INDCK-02	SAN JACINTO	232	236	137	157	148	148	236	236	270	278	192	212	199	199	33.761964	-116.883911	
INDCK-03	SAN JACINTO	232	256	149	157	148	148	236	236	250	270	204	204	219	235	33.761964	-116.883911	
INDCK-04	SAN JACINTO	232	232	145	149	148	148	236	236	270	274	192	204	235	235	33.761964	-116.883911	
INDCK-05	SAN JACINTO	232	232	137	137	128	148	236	244	262	266	192	204	199	219	33.761964	-116.883911	
INDCK-06	SAN JACINTO	232	232	149	157	148	148	236	236	270	270	192	192	195	199	33.761964	-116.883911	
INDCK-07	SAN JACINTO	256	256	137	137	184	216	236	236	270	270	192	192	199	219	33.761964	-116.883911	
MILL-1	SAN BERNARDINO	252	256	137	141	156	160	232	252	282	282	192	200	219	227	34.077789	-117.09945	
MILL-2	SAN BERNARDINO	252	264	137	141	160	160	232	252	274	282	216	216	203	207	34.077789	-117.09945	
MILL-3	SAN BERNARDINO	264	276	137	137	160	160	228	252	282	282	192	264	207	215	34.077789	-117.09945	
MILL-4	SAN BERNARDINO	236	276	137	137	140	172	232	232	274	282	200	220	199	215	34.077789	-117.09945	

MILL-5	SAN BERNARDINO	228	272	133	137	140	160	252	252	274	282	200	216	203	215	34.077789	-117.09945
PC-01	SAN BERNARDINO	252	252	133	137	140	188	228	228	238	238	192	196	223	227	34.108917	-117.15075
PC-02	SAN BERNARDINO	236	280	137	137	116	116	256	256	238	238	192	196	223	227	34.108917	-117.15075
PC-03	SAN BERNARDINO	228	264	137	141	180	192	248	248	238	238	204	220	227	227	34.108917	-117.15075
PC-04	SAN BERNARDINO	252	264	137	157	140	152	224	248	302	302	200	204	203	215	34.108917	-117.15075
PC-05	SAN BERNARDINO	252	276	141	157	140	160	256	256	278	286	204	224	223	223	34.108917	-117.15075
PC-06	SAN BERNARDINO	252	280	137	157	140	180	256	256	254	282	204	268	195	219	34.108917	-117.15075
PC-07	SAN BERNARDINO	264	280	137	137	140	180	232	232	274	258	220	268	203	227	34.108917	-117.15075
PC-08	SAN BERNARDINO	252	280	133	145	152	152	228	256	286	286	192	204	231	203	34.108917	-117.15075
PC-09	SAN BERNARDINO	252	280	137	137	152	192	232	232	254	274	196	204	203	227	34.108917	-117.15075
PC-10	SAN BERNARDINO	276	276	133	137	140	140	228	228	254	266	192	220	203	247	34.108917	-117.15075
PC-11	SAN BERNARDINO	228	268	137	141	168	168	244	244	278	282	200	204	227	277	34.108917	-117.15075
PC-12	SAN BERNARDINO	228	268	133	137	168	192	248	248	266	278	192	220	195	203	34.108917	-117.15075
PC-13	SAN BERNARDINO	236	280	137	145	136	184	248	248	246	274	200	200	215	227	34.108917	-117.15075
PC-14	SAN BERNARDINO	252	280	137	137	168	180	252	256	274	274	196	204	215	215	34.108917	-117.1507

PC-15	SAN BERNARDINO	252	276	141	141	168	168	228	256	290	290	200	200	175	175	34.108917	-117.15075
PC-16	SAN BERNARDINO	236	252	141	161	148	168	232	252	282	282	192	200	199	203	34.108917	-117.15075
PLNG-1	SAN BERNARDINO	236	276	137	141	140	184	248	248	266	290	200	204	203	203	34.108917	-117.15075
PLNG-2	SAN BERNARDINO	252	268	157	157	140	172	228	256	258	290	204	268	215	219	34.108917	-117.15075
PLNG-3	SAN BERNARDINO	276	276	137	145	140	184	244	244	266	286	192	204	219	227	34.108917	-117.15075
PLNG-4	SAN BERNARDINO	264	276	137	161	172	172	244	244	266	266	204	220	215	219	34.108917	-117.15075
PLUN-1	SAN BERNARDINO	276	276	157	161	152	172	256	256	266	266	196	224	203	219	34.108917	-117.15075
TWIN-01	SAN BERNARDINO	256	272	133	133	152	152	264	264	262	262	184	196	219	227	34.174722	-117.266667
TWIN-02	SAN BERNARDINO	224	224	133	137	192	192	232	264	262	290	184	200	211	223	34.174722	-117.266667
TWIN-03	SAN BERNARDINO	224	284	153	153	192	196	264	268	262	262	200	212	211	211	34.174722	-117.266667
TWIN-04	SAN BERNARDINO	224	264	133	137	160	160	256	268	266	270	200	212	211	223	34.174722	-117.266667
TWIN-05	SAN BERNARDINO	264	264	133	133	156	168	228	264	290	290	184	200	211	219	34.174722	-117.266667
TWIN-06	SAN BERNARDINO	260	292	133	133	156	168	256	268	266	290	200	200	211	227	34.174722	-117.266667
TWIN-07	SAN BERNARDINO	260	284	121	133	160	160	260	260	274	282	200	200	219	219	34.174722	-117.266667

TWIN-08	SAN BERNARDINO	252	252	133	133	192	192	252	264	262	282	200	200	211	211	34.174722	-117.266667
TWIN-09	SAN BERNARDINO	256	268	137	141	156	192	264	264	254	274	196	200	211	227	34.174722	-117.266667
TWIN-10	SAN BERNARDINO	224	264	137	145	156	168	264	264	262	274	200	200	211	211	34.174722	-117.266667
TWIN-11	SAN BERNARDINO	256	292	133	133	156	192	252	252	278	278	200	200	203	215	34.174722	-117.266667
TWIN-12	SAN BERNARDINO	252	280	133	137	156	192	236	264	270	270	200	200	211	227	34.174722	-117.266667
CC-01	SAN BERNARDINO	236	244	149	161	152	160	228	228	270	270	184	184	235	235	34.292526	-117.306135
CC-02	SAN BERNARDINO	244	252	137	145	140	164	260	260	278	282	184	200	227	235	34.292526	-117.306135
CC-03	SAN BERNARDINO	236	236	137	149	140	156	244	256	274	274	188	196	227	227	34.292526	-117.306135
CC-04	SAN BERNARDINO	244	276	145	157	156	172	244	244	274	278	196	196	231	235	34.292526	-117.306135
CC-05	SAN BERNARDINO	244	276	141	165	164	172	256	256	270	274	184	188	219	227	34.292526	-117.306135
CC-06	SAN BERNARDINO	232	272	129	149	164	168	244	260	274	294	208	220	227	227	34.292526	-117.306135
CC-07	SAN BERNARDINO	244	244	141	145	152	152	244	244	274	274	184	188	227	231	34.292526	-117.306135
CC-08	SAN BERNARDINO	232	272	141	149	168	168	244	244	258	274	196	204	227	227	34.292526	-117.306135
CC-09	SAN BERNARDINO	240	276	137	153	168	168	244	244	274	278	200	208	227	227	34.292526	-117.306135
CC-10	SAN BERNARDINO	240	276	141	153	156	200	244	256	274	274	204	220	227	227	34.292526	-117.3061

																	35
CC-11	SAN BERNARDINO	244	244	125	133	156	168	240	248	278	278	188	220	227	227	34.292526	-117.306135
CCWF-01	SAN BERNARDINO	236	244	149	161	164	208	228	228	274	274	188	192	227	235	34.292526	-117.306135
FORK-1	SAN BERNARDINO	224	224	129	137	164	164	236	256	282	294	188	196	235	247	34.292526	-117.306135
FORK-2	SAN BERNARDINO	224	244	137	145	132	172	208	208	254	294	200	212	199	227	34.292526	-117.306135
CAJO-01	SAN GABRIEL	264	276	141	141	172	172	232	252	254	270	196	264	199	231	34.232811	-117.427183
CAJO-02	SAN GABRIEL	252	252	137	137	156	156	208	224	266	282	196	236	203	231	34.232811	-117.427183
CAJO-03	SAN GABRIEL	268	280	125	157	168	172	228	228	254	270	196	196	215	247	34.232811	-117.427183
CAJO-04	SAN GABRIEL	264	268	133	153	216	220	240	256	254	274	196	212	191	227	34.232811	-117.427183
CAJO-05	SAN GABRIEL	276	276	125	157	168	172	228	228	254	270	196	196	215	247	34.232811	-117.427183
CAJO-06	SAN GABRIEL	224	264	133	145	156	172	232	232	254	270	212	212	227	247	34.232811	-117.427183
CAJO-07	SAN GABRIEL	220	252	145	149	176	176	208	228	254	270	204	208	211	235	34.232811	-117.427183
CAJO-08	SAN GABRIEL	224	224	141	145	132	172	224	240	250	254	188	196	223	231	34.232811	-117.427183
CAJO-09	SAN GABRIEL	280	280	141	141	132	216	228	232	254	274	208	208	199	227	34.232811	-117.427183
CAJO-10	SAN GABRIEL	272	276	133	141	176	176	232	244	254	266	212	216	191	227	34.232811	-117.427183

CAJO-11	SAN GABRIEL	260	264	145	145	164	164	228	240	274	294	196	204	199	227	34.232811	-117.427183
CAJO-12	SAN GABRIEL	264	280	133	145	132	176	228	244	266	274	196	212	227	239	34.232811	-117.427183
CAJO-13	SAN GABRIEL	252	264	141	145	168	216	212	228	250	270	212	216	215	227	34.232811	-117.427183
CAJO-14	SAN GABRIEL	224	248	133	145	164	180	228	232	270	250	204	212	199	211	34.232811	-117.427183
CAJO-15	SAN GABRIEL	252	272	137	141	176	176	236	244	270	274	192	196	227	239	34.232811	-117.427183
CAJO-16	SAN GABRIEL	248	268	129	141	132	168	244	244	266	270	188	216	247	247	34.232811	-117.427183
CAJO-17	SAN GABRIEL	248	248	141	141	168	168	228	236	274	278	188	188	199	227	34.232811	-117.427183
CAJW-1	SAN GABRIEL	224	264	133	141	176	176	232	232	254	266	196	212	199	227	34.232811	-117.427183
CAJW-2	SAN GABRIEL	220	272	133	145	172	212	244	244	254	270	188	196	227	235	34.232811	-117.427183
CAJW-3	SAN GABRIEL	248	276	133	141	176	180	232	236	266	278	196	212	191	227	34.232811	-117.427183
CAJW-4	SAN GABRIEL	248	276	133	141	184	184	208	236	270	270	196	212	199	227	34.232811	-117.427183
LCKC-01	SAN GABRIEL	220	276	141	141	156	156	236	236	254	270	188	196	227	247	34.228583	-117.473517
LCKC-02	SAN GABRIEL	252	276	141	141	156	176	232	232	270	274	264	264	247	247	34.228583	-117.473517
LCKC-03	SAN GABRIEL	220	264	141	141	156	216	228	228	254	254	212	212	211	247	34.228583	-117.473517
LCKC-04	SAN GABRIEL	264	276	141	141	156	156	228	236	254	254	188	196	211	211	34.228583	-117.4735

LCKC-05	SAN GABRIEL	264	268	141	141	176	176	228	228	254	270	188	188	211	227	34.228583	-117.473517
LCKC-06	SAN GABRIEL	264	276	141	141	156	156	228	228	254	274	208	260	227	247	34.228583	-117.473517
LCKC-07	SAN GABRIEL	220	276	141	141	148	152	228	232	254	274	188	188	227	227	34.228583	-117.473517
LCKC-08	SAN GABRIEL	220	276	145	145	212	212	236	236	254	254	188	188	227	227	34.228583	-117.473517
LCKC-09	SAN GABRIEL	220	276	141	141	156	216	228	236	254	274	212	264	227	227	34.228583	-117.473517
LCKC-10	SAN GABRIEL	220	276	141	141	156	216	228	228	254	254	196	212	227	247	34.228583	-117.473517
LCK-11	SAN GABRIEL	220	276	141	141	176	216	228	232	254	254	188	188	211	227	34.228583	-117.473517
LCK-12	SAN GABRIEL	220	252	141	141	156	156	228	232	254	254	188	188	227	227	34.228583	-117.473517
LCK-13	SAN GABRIEL	220	268	137	141	156	216	228	236	270	274	188	196	247	247	34.228583	-117.473517
LCK-14	SAN GABRIEL	220	276	137	141	156	176	228	232	254	254	196	208	211	227	34.228583	-117.473517
LCK-15	SAN GABRIEL	252	260	141	145	172	176	228	236	254	254	196	256	211	227	34.228583	-117.473517
LCK-16	SAN GABRIEL	220	276	137	141	156	216	228	228	254	270	196	196	211	211	34.228583	-117.473517
LCK-17	SAN GABRIEL	220	276	137	141	156	216	232	232	254	270	264	264	211	227	34.228583	-117.473517
LCK-18	SAN GABRIEL	252	264	141	141	156	156	228	228	270	270	212	264	227	227	34.228583	-117.473517

LCK-19	SAN GABRIEL	252	260	145	145	156	156	228	236	254	270	188	188	227	277	34.228583	-117.473517
LCK-20	SAN GABRIEL	220	276	141	141	156	176	232	236	254	254	196	212	227	247	34.228583	-117.473517
LCK-21	SAN GABRIEL	260	276	141	141	156	156	228	236	254	254	264	264	211	227	34.228583	-117.473517
LCK-22	SAN GABRIEL	264	276	137	141	156	176	220	232	254	274	196	196	211	227	34.228583	-117.473517
LCK-23	SAN GABRIEL	220	264	141	141	156	156	228	232	254	274	188	212	227	247	34.228583	-117.473517
LCK-24	SAN GABRIEL	220	268	141	141	156	156	228	232	274	274	196	196	211	227	34.228583	-117.473517
LCK-25	SAN GABRIEL	220	220	133	141	156	156	232	232	254	282	188	212	247	247	34.228583	-117.473517
LCK-26	SAN GABRIEL	264	276	141	141	156	176	232	236	254	274	188	188	227	227	34.228583	-117.473517
LCK-27	SAN GABRIEL	264	276	141	141	152	152	224	228	254	270	196	212	227	247	34.228583	-117.473517
LCK-28	SAN GABRIEL	220	280	141	141	172	172	220	228	254	274	196	260	211	247	34.228583	-117.473517
CATCR-01	SAN GABRIEL	244	284	125	137	160	164	244	248	258	262	212	216	219	231	34.23	-117.778832
CATCR-02	SAN GABRIEL	228	252	125	161	152	156	252	268	278	278	228	228	203	207	34.23	-117.778832
CATCR-03	SAN GABRIEL	252	264	125	141	160	172	232	256	270	274	224	228	231	239	34.23	-117.778832
EFSGR-01	SAN GABRIEL	240	244	137	141	160	164	276	260	266	298	220	228	227	239	34.258056	-118.103611
EFSGR-02	SAN GABRIEL	244	252	137	145	160	164	276	340	278	298	224	232	199	231	34.258056	-118.1036

NFSGR-01	SAN GABRIEL	264	272	133	141	152	172	232	256	266	294	196	208	187	235	34.258056	-118.103611
NFSGR-02	SAN GABRIEL	240	264	129	145	156	172	252	264	266	294	196	232	235	255	34.258056	-118.103611
FISH-1	SAN GABRIEL	252	260	133	133	152	168	248	252	266	270	212	228	203	231	34.169372	-117.925931
FISH-2	SAN GABRIEL	276	288	133	133	152	172	268	268	262	274	224	244	223	231	34.169372	-117.925931
FISH-3	SAN GABRIEL	220	264	133	137	152	152	220	220	270	270	224	244	203	203	34.169372	-117.925931
FISH-4	SAN GABRIEL	260	284	133	133	152	176	248	248	266	266	208	212	203	203	34.169372	-117.925931
FISH-5	SAN GABRIEL	260	260	129	133	152	172	248	248	262	262	240	244	215	219	34.169372	-117.925931
WFSGR-01	SAN GABRIEL	252	260	137	153	160	164	260	260	278	278	192	240	223	235	34.258056	-118.103611
HAIN-1	SAN GABRIEL	260	260	117	145	148	148	232	276	270	274	216	236	203	203	34.236788	-118.282027
HAIN-2	SAN GABRIEL	260	268	145	145	132	152	232	276	270	274	204	236	219	219	34.236788	-118.282027

STRUCTURE PARAMETER SET: SIMULATION CONFIGURATION

Parameter Set: 100K_1MIL

Running Length

Length of Burnin Period: 100000

Number of MCMC Reps after Burnin: 1000000

Ancestry Model Info

Use Admixture Model

* Use Sampling Location Information

* Use Population IDs as Sampling Location Information

* Infer Alpha

* Initial Value of ALPHA (Dirichlet Parameter for Degree of Admixture): 1.0

* Use Same Alpha for all Populations

* Use a Uniform Prior for Alpha

** Maximum Value for Alpha: 10.0

** SD of Proposal for Updating Alpha: 0.025

Frequency Model Info

Allele Frequencies are Correlated among Pops

* Assume Different Values of Fst for Different Subpopulations

* Prior Mean of Fst for Pops: 0.01

* Prior SD of Fst for Pops: 0.05

* Use Constant Lambda (Allele Frequencies Parameter)

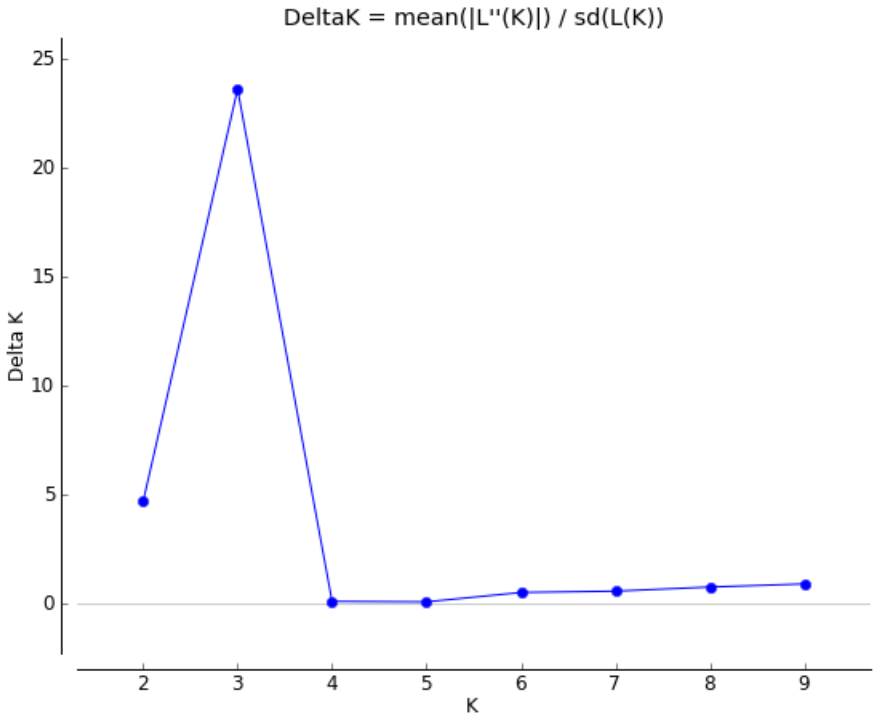
* Value of Lambda: 1.0

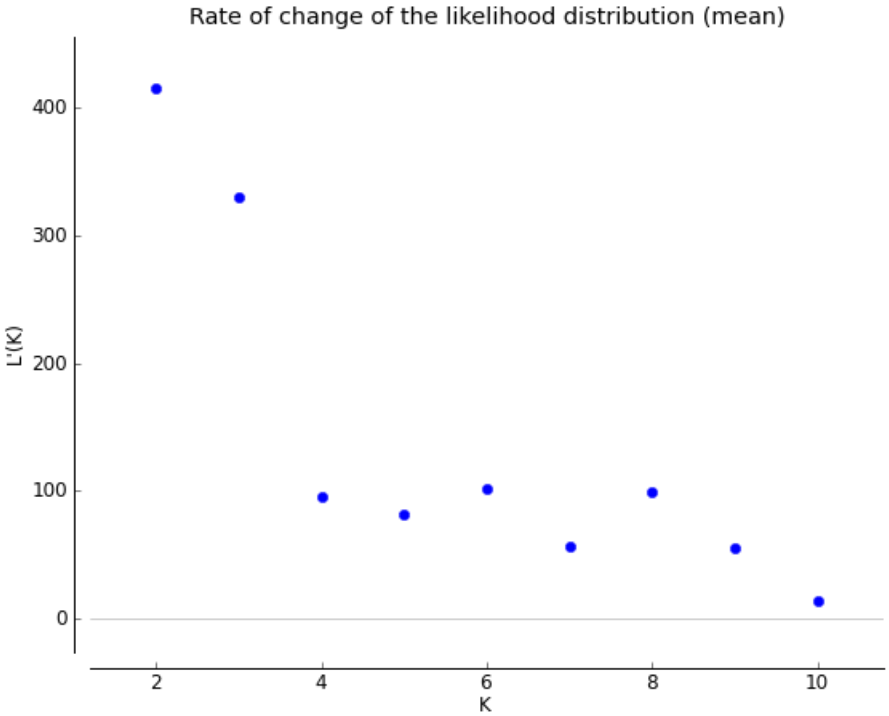
Advanced Options

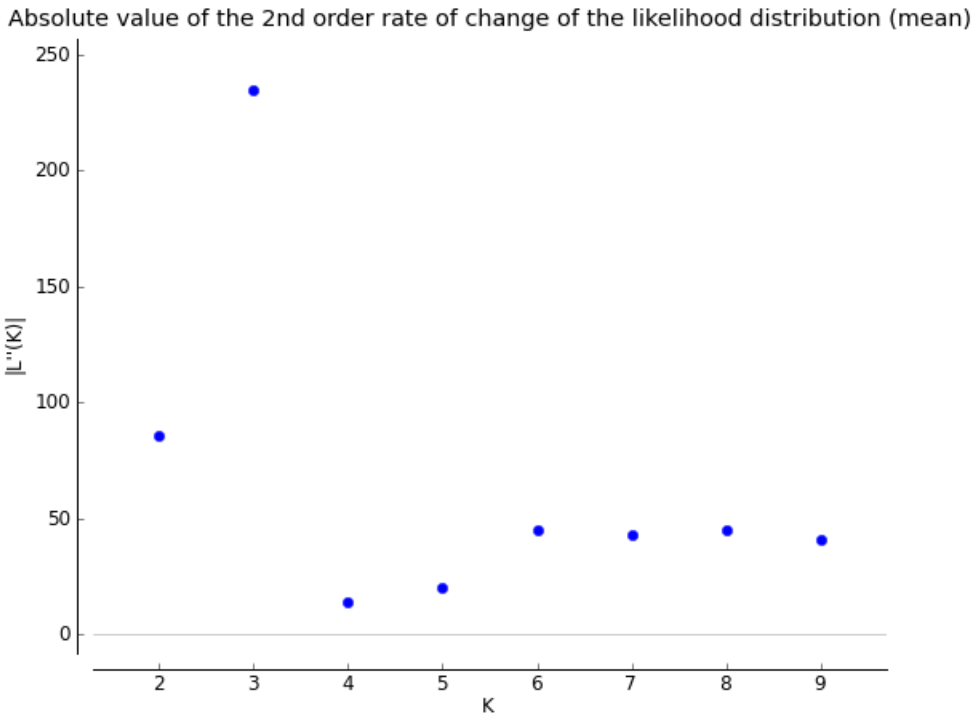
Estimate the Probability of the Data Under the Model

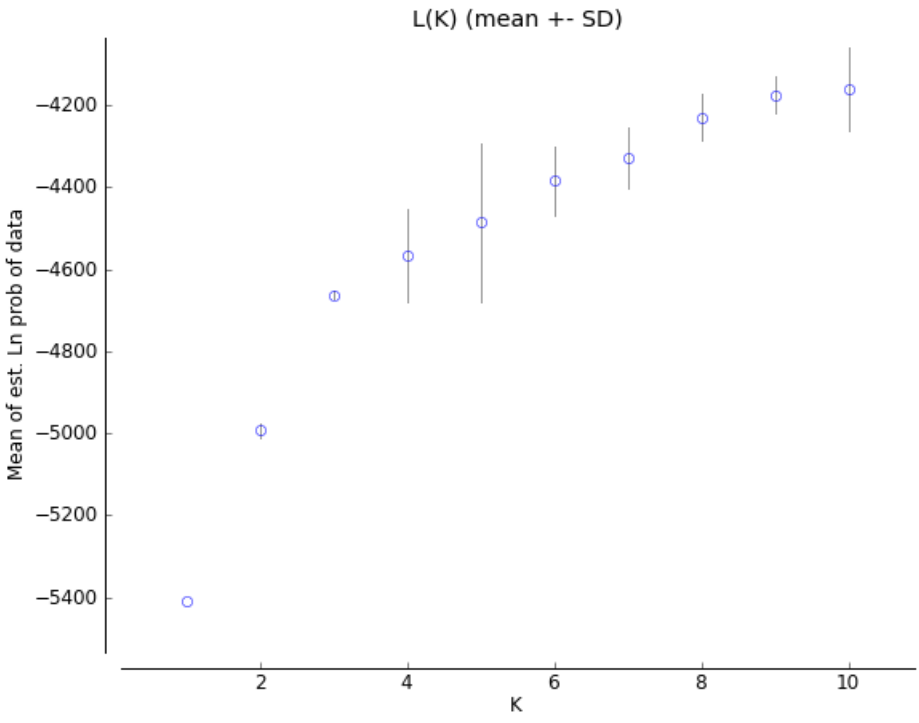
Frequency of Metropolis update for Q: 10

STRUCTURE RESULTS K=3 (CALIFORNIA POPULATIONS)

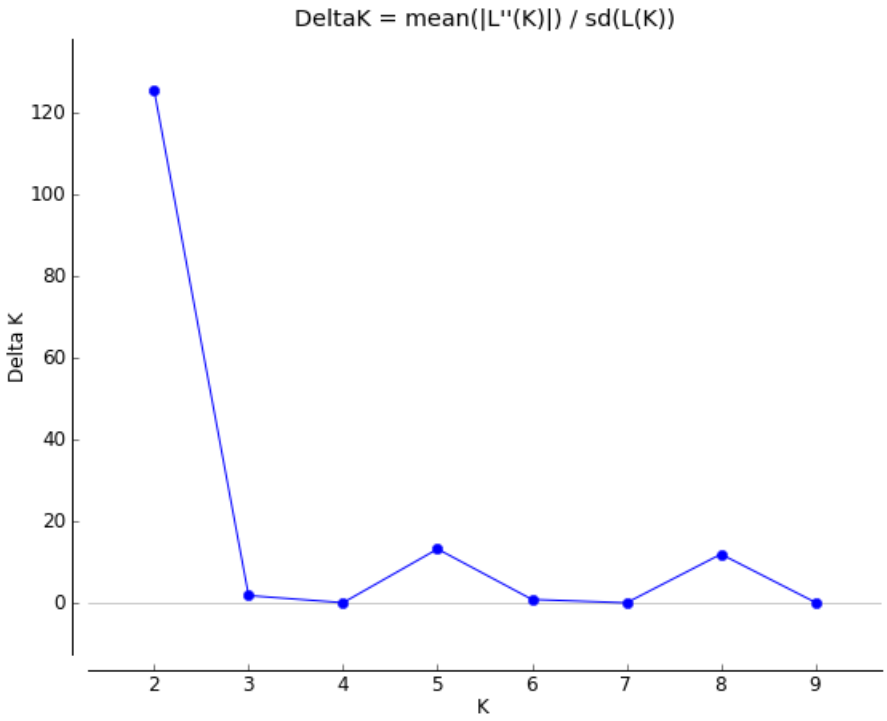


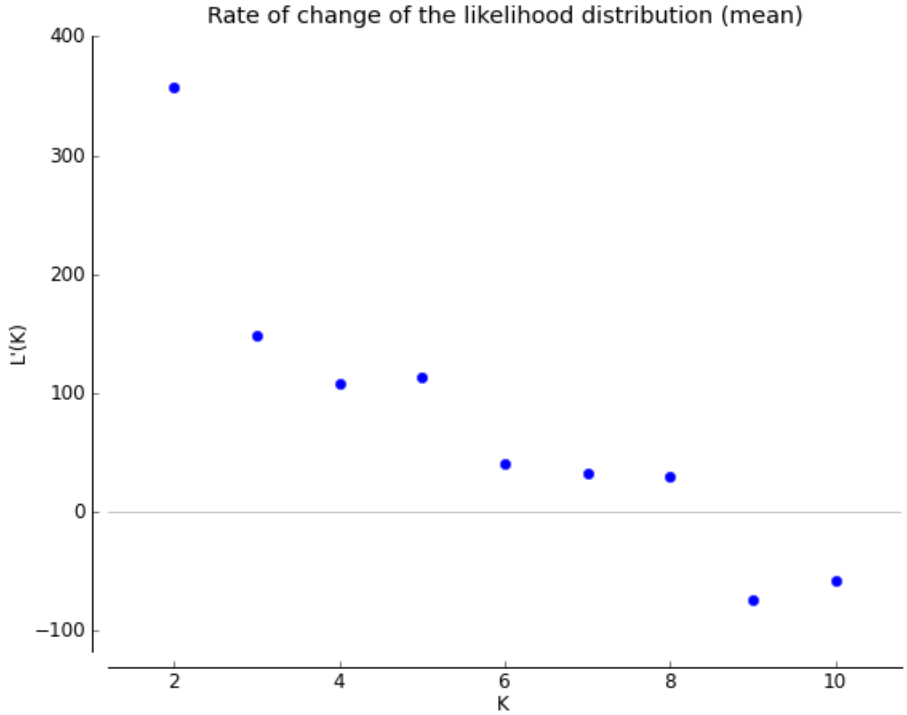


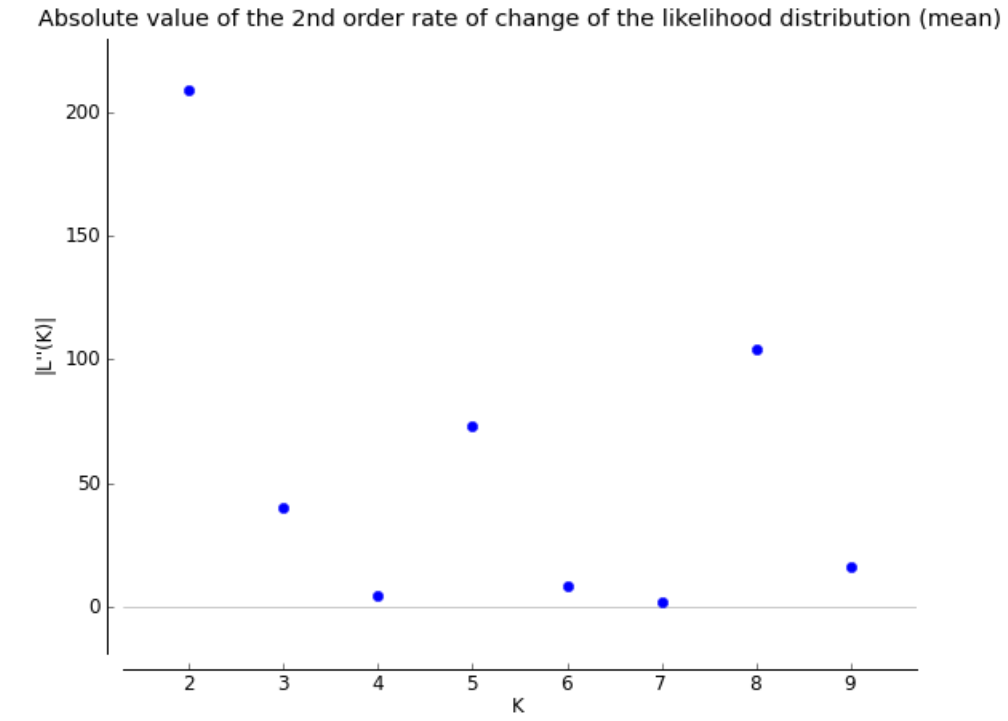

Absolute value of the 2nd order rate of change of the likelihood distribution (mean)
|L''(K)|
250
200
150
100
50
0
2
3
4
5
6
7
8
9
K

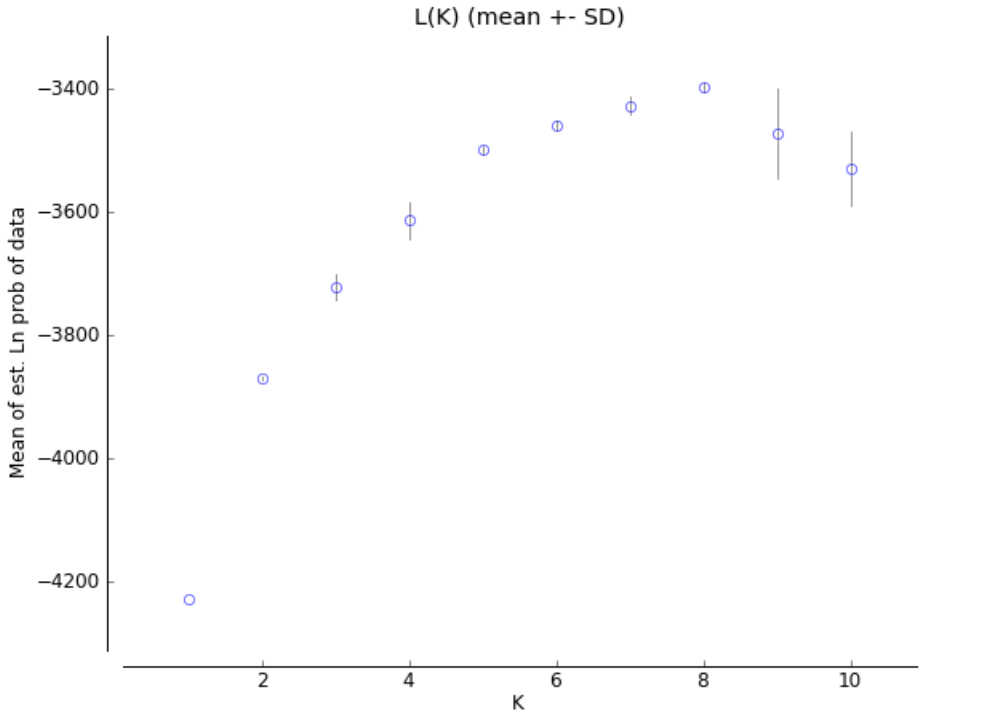

L(K) (mean +- SD)
Mean of est. Ln prob of data
−4200
−4400
−4600
−4800
−5000
−5200
−5400
2
4
6
8
10
K

STRUCTURE RESULTS K=2 (SOUTHERN CALIFORNIA POPULATIONS)






Absolute value of the 2nd order rate of change of the likelihood distribution (mean)
|L''(K)|
200
150
100
50
0
2
3
4
5
6
7
8
9
K


L(K) (mean +- SD)
Mean of est. Ln prob of data
−3400
−3600
−3800
−4000
−4200
2
4
6
8
10
K

CLUMPP K=3 INDIVIDUAL PARAMETER FILE

```
# This is the file that sets the parameters for the program
CLUMPP, version 1.1

# Everything after "#" will be ignored by the program. Parameters
are:
# K, C, R, M, W, GREEDY_OPTION, REPEATS, PERMUTATIONFILE,
PRINT_PERMUTED_DATA,
# PERMUTED_DATAFILE, PRINT_EVERY_PERM and EVERY_PERMFILE.
# All parameter names shall be followed by at least one blank
space and then
# the parameter-value.



# --------------- Main parameters ------------------------------
--------------

DATATYPE 0                          # The type of data to be
read in.
                                    # 0 = individual data in
the file
                                    # specified by INDFILE, 1 =
population
                                    # data in the file
specified by
                                    # POPFILE.

INDFILE K3.indfile              # The name of the individual
datafile.
                                    # Required if DATATYPE = 0.

POPFILE k3.popfile                  # The name of the
population datafile.
                                    # Required if DATATYPE = 1.

OUTFILE k3.outfile                  # The average cluster
membership
                                    # coefficients across the
permuted runs
                                    # are printed here.

MISCFILE k3.miscfile            # The parameters used and a
summary of
                                    # the results are printed
here.
```

```
K 3                                     # Number of clusters.

C 146                                   # Number of individuals or
populations.

R 25                                    # Number of runs.

M 2                                     # Method to be used (1 =
FullSearch,
                                        # 2 = Greedy, 3 =
LargeKGreedy).

W 1                                     # Weight by the number of
individuals
                                        # in each population as
specified in
                                        # the datafile (1 if yes, 0
if no).

S 2                                     # Pairwise matrix
similarity statistic
                                        # to be used. 1 = G, 2 =
G'.


# - Additional options for the Greedy and LargeKGreedy algorithm
(M = 2 or 3) -

GREEDY_OPTION 2                         # 1 = All possible input
orders,
                                        # 2 = random input orders,
                                        # 3 = pre-specified input
orders.

REPEATS 1000                            # If GREEDY_OPTION = 2,
then REPEATS
                                        # determines the number of
random input
                                        # orders to be tested. If
GREEDY_OPTION
                                        # = 3, then REPEATS is the
number of
                                        # input orders in
PERMUTATIONFILE.

PERMUTATIONFILE arabid.permutationfile      # The permutations of
the runs in
                                        # PERMUTATIONFILE will be
used, if
```

```
                                        # GREEDY_OPTION = 3.


# --------------- Optional outputs ----------------------------
--------------

PRINT_PERMUTED_DATA 1                   # Print the permuted data
(clusters) in
                                        # INDFILE or POPFILE to
                                        # PERMUTED_DATAFILE (0 =
don't print,
                                        # 1 = print into one file,
2 = print
                                        # into separate files for
each run).

PERMUTED_DATAFILE arabid.perm_datafile# The permuted data
(clusters) will be
                                        # printed to this file (if
                                        # PRINT_PERMUTED_DATA = 2,
several
                                        # files with the extensions
"_1" to
                                        # "_R" will be created).

PRINT_EVERY_PERM 0                      # Print every tested
permutation of the
                                        # runs and the
corresponding value of
                                        # SSC to a file specified
by
                                        # EVERY_PERMFILE (0 = don't
print,
                                        # 1 = print).
                                        # Note that printing may
result in a
                                        # very large file.

EVERY_PERMFILE arabid.every_permfile  # Every tested permutation
of the runs
                                        # and the corresponding SSC
will be
                                        # printed here.

PRINT_RANDOM_INPUTORDER 0               # Print random input orders
of runs to
                                        # RANDOM_INPUTORDER (0 =
don't print,
```

```
                                                # 1 = print). This
option is only
                                                # available if
GREEDY_OPTION = 2.

RANDOM_INPUTORDERFILE arabid.random_inputorderfile # Every random
input order
                                                # of the runs
(generated by CLUMPP if
                                                # GREEDY_OPTION = 2)
will be printed
                                                # here.


# --------------- Advanced options ------------------------------
--------------

OVERRIDE_WARNINGS 0                     # This option allows the
user to
                                        # override non-crucial
warnings from
                                        # the program (0 allow
warnings, 1 do
                                        # not issue non-crucial
warnings).

ORDER_BY_RUN 1                          # Permute the clusters of
the output
                                        # files by the specified
run. (0 to
                                        # not specify a run, 1 to R
specifies
                                        # a run in the INDFILE or
POPFILE).


# --------------- Additional comments --------------------------
--------------

# The term ''permutation'' is used in two different contexts,
permutations of
# membership coefficients, or clusters, and permutations of runs.

# For example, if the datafile has has data A B C D E (each
letter indicates a
# column corresponding to a cluster), then permutation 3 2 5 1 4
of the
# clusters means C B E A D.
```

```
# Permutation 4 1 2 3 of runs 1-4 would mean start with run 4,
then run 1, then
# run 2, and then run 3.


# --------------- Command line arguments ------------------------
--------------

# -i INDFILE
# -p POPFILE
# -o OUTFILE
# -j MISCFILE
# -k K
# -c C
# -r R
# -m M
# -w W
# -s S
# ----------------------------------------------------------------
--------------
```

CLUMPP K=3 POPULATION PARAMETER FILE

```
# This is the file that sets the parameters for the program
CLUMPP, version 1.1

# Everything after "#" will be ignored by the program. Parameters
are:
# K, C, R, M, W, GREEDY_OPTION, REPEATS, PERMUTATIONFILE,
PRINT_PERMUTED_DATA,
# PERMUTED_DATAFILE, PRINT_EVERY_PERM and EVERY_PERMFILE.
# All parameter names shall be followed by at least one blank
space and then
# the parameter-value.



# --------------- Main parameters ------------------------------
--------------

DATATYPE 1                          # The type of data to be
read in.
                                    # 0 = individual data in
the file
                                    # specified by INDFILE, 1 =
population
                                    # data in the file
specified by
                                    # POPFILE.

INDFILE K3.popfile              # The name of the individual
datafile.
                                    # Required if DATATYPE = 0.

POPFILE k3.popfile                  # The name of the
population datafile.
                                    # Required if DATATYPE = 1.

OUTFILE k3.outfile                  # The average cluster
membership
                                    # coefficients across the
permuted runs
                                    # are printed here.

MISCFILE k3.miscfile            # The parameters used and a
summary of
                                    # the results are printed
here.
```

```
K 3                                     # Number of clusters.

C 3                                     # Number of individuals or
populations.

R 25                                    # Number of runs.

M 2                                     # Method to be used (1 =
FullSearch,
                                        # 2 = Greedy, 3 =
LargeKGreedy).

W 1                                     # Weight by the number of
individuals
                                        # in each population as
specified in
                                        # the datafile (1 if yes, 0
if no).

S 2                                     # Pairwise matrix
similarity statistic
                                        # to be used. 1 = G, 2 =
G'.


# - Additional options for the Greedy and LargeKGreedy algorithm
(M = 2 or 3) -

GREEDY_OPTION 2                         # 1 = All possible input
orders,
                                        # 2 = random input orders,
                                        # 3 = pre-specified input
orders.

REPEATS 1000                            # If GREEDY_OPTION = 2,
then REPEATS
                                        # determines the number of
random input
                                        # orders to be tested. If
GREEDY_OPTION
                                        # = 3, then REPEATS is the
number of
                                        # input orders in
PERMUTATIONFILE.

PERMUTATIONFILE arabid.permutationfile      # The permutations of
the runs in
                                        # PERMUTATIONFILE will be
used, if
```

```
                                        # GREEDY_OPTION = 3.


# --------------- Optional outputs ----------------------------
--------------

PRINT_PERMUTED_DATA 1                   # Print the permuted data
(clusters) in
                                        # INDFILE or POPFILE to
                                        # PERMUTED_DATAFILE (0 =
don't print,
                                        # 1 = print into one file,
2 = print
                                        # into separate files for
each run).

PERMUTED_DATAFILE arabid.perm_datafile# The permuted data
(clusters) will be
                                        # printed to this file (if
                                        # PRINT_PERMUTED_DATA = 2,
several
                                        # files with the extensions
"_1" to
                                        # "_R" will be created).

PRINT_EVERY_PERM 0                      # Print every tested
permutation of the
                                        # runs and the
corresponding value of
                                        # SSC to a file specified
by
                                        # EVERY_PERMFILE (0 = don't
print,
                                        # 1 = print).
                                        # Note that printing may
result in a
                                        # very large file.

EVERY_PERMFILE arabid.every_permfile  # Every tested permutation
of the runs
                                        # and the corresponding SSC
will be
                                        # printed here.

PRINT_RANDOM_INPUTORDER 0               # Print random input orders
of runs to
                                        # RANDOM_INPUTORDER (0 =
don't print,
```

```
                                                # 1 = print). This
option is only
                                                # available if
GREEDY_OPTION = 2.

RANDOM_INPUTORDERFILE arabid.random_inputorderfile # Every random
input order
                                                # of the runs
(generated by CLUMPP if
                                                # GREEDY_OPTION = 2)
will be printed
                                                # here.


# --------------- Advanced options ------------------------------
--------------

OVERRIDE_WARNINGS 0                     # This option allows the
user to
                                        # override non-crucial
warnings from
                                        # the program (0 allow
warnings, 1 do
                                        # not issue non-crucial
warnings).

ORDER_BY_RUN 1                          # Permute the clusters of
the output
                                        # files by the specified
run. (0 to
                                        # not specify a run, 1 to R
specifies
                                        # a run in the INDFILE or
POPFILE).


# --------------- Additional comments --------------------------
--------------

# The term ''permutation'' is used in two different contexts,
permutations of
# membership coefficients, or clusters, and permutations of runs.

# For example, if the datafile has has data A B C D E (each
letter indicates a
# column corresponding to a cluster), then permutation 3 2 5 1 4
of the
# clusters means C B E A D.
```

```
# Permutation 4 1 2 3 of runs 1-4 would mean start with run 4,
then run 1, then
# run 2, and then run 3.


# --------------- Command line arguments ------------------------
--------------

# -i INDFILE
# -p POPFILE
# -o OUTFILE
# -j MISCFILE
# -k K
# -c C
# -r R
# -m M
# -w W
# -s S
# ---------------------------------------------------------------
--------------
```

CLUMPP K=2 INDIVIDUAL PARAMETER FILE

```
# This is the file that sets the parameters for the program
CLUMPP, version 1.1

# Everything after "#" will be ignored by the program. Parameters
are:
# K, C, R, M, W, GREEDY_OPTION, REPEATS, PERMUTATIONFILE,
PRINT_PERMUTED_DATA,
# PERMUTED_DATAFILE, PRINT_EVERY_PERM and EVERY_PERMFILE.
# All parameter names shall be followed by at least one blank
space and then
# the parameter-value.



# --------------- Main parameters ------------------------------
--------------

DATATYPE 0                              # The type of data to be
read in.
                                        # 0 = individual data in
the file
                                        # specified by INDFILE, 1 =
population
                                        # data in the file
specified by
                                        # POPFILE.

INDFILE k2.indfile                  # The name of the individual
datafile.
                                        # Required if DATATYPE = 0.

POPFILE k2.popfile                      # The name of the
population datafile.
                                        # Required if DATATYPE = 1.

OUTFILE k2.outfile                      # The average cluster
membership
                                        # coefficients across the
permuted runs
                                        # are printed here.

MISCFILE k2.miscfile                # The parameters used and a
summary of
                                        # the results are printed
here.
```

```
K 2                                     # Number of clusters.

C 123                                   # Number of individuals or
populations.

R 15                                    # Number of runs.

M 2                                     # Method to be used (1 =
FullSearch,
                                        # 2 = Greedy, 3 =
LargeKGreedy).

W 1                                     # Weight by the number of
individuals
                                        # in each population as
specified in
                                        # the datafile (1 if yes, 0
if no).

S 2                                     # Pairwise matrix
similarity statistic
                                        # to be used. 1 = G, 2 =
G'.


# - Additional options for the Greedy and LargeKGreedy algorithm
(M = 2 or 3) -

GREEDY_OPTION 2                         # 1 = All possible input
orders,
                                        # 2 = random input orders,
                                        # 3 = pre-specified input
orders.

REPEATS 1000                            # If GREEDY_OPTION = 2,
then REPEATS
                                        # determines the number of
random input
                                        # orders to be tested. If
GREEDY_OPTION
                                        # = 3, then REPEATS is the
number of
                                        # input orders in
PERMUTATIONFILE.

PERMUTATIONFILE arabid.permutationfile      # The permutations of
the runs in
                                        # PERMUTATIONFILE will be
used, if
```

```
                                        # GREEDY_OPTION = 3.


# --------------- Optional outputs ------------------------------
--------------

PRINT_PERMUTED_DATA 1                   # Print the permuted data
(clusters) in
                                        # INDFILE or POPFILE to
                                        # PERMUTED_DATAFILE (0 =
don't print,
                                        # 1 = print into one file,
2 = print
                                        # into separate files for
each run).

PERMUTED_DATAFILE arabid.perm_datafile# The permuted data
(clusters) will be
                                        # printed to this file (if
                                        # PRINT_PERMUTED_DATA = 2,
several
                                        # files with the extensions
"_1" to
                                        # "_R" will be created).

PRINT_EVERY_PERM 0                      # Print every tested
permutation of the
                                        # runs and the
corresponding value of
                                        # SSC to a file specified
by
                                        # EVERY_PERMFILE (0 = don't
print,
                                        # 1 = print).
                                        # Note that printing may
result in a
                                        # very large file.

EVERY_PERMFILE arabid.every_permfile  # Every tested permutation
of the runs
                                        # and the corresponding SSC
will be
                                        # printed here.

PRINT_RANDOM_INPUTORDER 0               # Print random input orders
of runs to
                                        # RANDOM_INPUTORDER (0 =
don't print,
```

```
                                                # 1 = print). This
option is only
                                                # available if
GREEDY_OPTION = 2.

RANDOM_INPUTORDERFILE arabid.random_inputorderfile # Every random
input order
                                                # of the runs
(generated by CLUMPP if
                                                # GREEDY_OPTION = 2)
will be printed
                                                # here.


# --------------- Advanced options ------------------------------
--------------

OVERRIDE_WARNINGS 0                     # This option allows the
user to
                                        # override non-crucial
warnings from
                                        # the program (0 allow
warnings, 1 do
                                        # not issue non-crucial
warnings).

ORDER_BY_RUN 1                          # Permute the clusters of
the output
                                        # files by the specified
run. (0 to
                                        # not specify a run, 1 to R
specifies
                                        # a run in the INDFILE or
POPFILE).


# --------------- Additional comments --------------------------
--------------

# The term ''permutation'' is used in two different contexts,
permutations of
# membership coefficients, or clusters, and permutations of runs.

# For example, if the datafile has has data A B C D E (each
letter indicates a
# column corresponding to a cluster), then permutation 3 2 5 1 4
of the
# clusters means C B E A D.
```

```
# Permutation 4 1 2 3 of runs 1-4 would mean start with run 4,
then run 1, then
# run 2, and then run 3.


# --------------- Command line arguments ------------------------
--------------

# -i INDFILE
# -p POPFILE
# -o OUTFILE
# -j MISCFILE
# -k K
# -c C
# -r R
# -m M
# -w W
# -s S
# ----------------------------------------------------------------
--------------
```

CLUMPP K=2 POPULATION

```
# This is the file that sets the parameters for the program
CLUMPP, version 1.1

# Everything after "#" will be ignored by the program. Parameters
are:
# K, C, R, M, W, GREEDY_OPTION, REPEATS, PERMUTATIONFILE,
PRINT_PERMUTED_DATA,
# PERMUTED_DATAFILE, PRINT_EVERY_PERM and EVERY_PERMFILE.
# All parameter names shall be followed by at least one blank
space and then
# the parameter-value.



# --------------- Main parameters ------------------------------
--------------

DATATYPE 1                              # The type of data to be
read in.
                                        # 0 = individual data in
the file
                                        # specified by INDFILE, 1 =
population
                                        # data in the file
specified by
                                        # POPFILE.

INDFILE k2.indfile                      # The name of the
individual datafile.
                                        # Required if DATATYPE = 0.

POPFILE k2.popfile                      # The name of the
population datafile.
                                        # Required if DATATYPE = 1.

OUTFILE k2.outfile                      # The average cluster
membership
                                        # coefficients across the
permuted runs
                                        # are printed here.

MISCFILE k2.miscfile             # The parameters used and a
summary of
                                        # the results are printed
here.

K 2                                     # Number of clusters.
```

```
C 13                                    # Number of individuals or
populations.

R 15                                    # Number of runs.

M 2                                     # Method to be used (1 =
FullSearch,
                                        # 2 = Greedy, 3 =
LargeKGreedy).

W 1                                     # Weight by the number of
individuals
                                        # in each population as
specified in
                                        # the datafile (1 if yes, 0
if no).

S 2                                     # Pairwise matrix
similarity statistic
                                        # to be used. 1 = G, 2 =
G'.


# - Additional options for the Greedy and LargeKGreedy algorithm
(M = 2 or 3) -

GREEDY_OPTION 2                         # 1 = All possible input
orders,
                                        # 2 = random input orders,
                                        # 3 = pre-specified input
orders.

REPEATS 1000                            # If GREEDY_OPTION = 2,
then REPEATS
                                        # determines the number of
random input
                                        # orders to be tested. If
GREEDY_OPTION
                                        # = 3, then REPEATS is the
number of
                                        # input orders in
PERMUTATIONFILE.

PERMUTATIONFILE arabid.permutationfile      # The permutations of
the runs in
                                        # PERMUTATIONFILE will be
used, if
                                        # GREEDY_OPTION = 3.
```

```
# --------------- Optional outputs -----------------------------
--------------

PRINT_PERMUTED_DATA 1                  # Print the permuted data
(clusters) in
                                       # INDFILE or POPFILE to
                                       # PERMUTED_DATAFILE (0 =
don't print,
                                       # 1 = print into one file,
2 = print
                                       # into separate files for
each run).

PERMUTED_DATAFILE arabid.perm_datafile# The permuted data
(clusters) will be
                                       # printed to this file (if
                                       # PRINT_PERMUTED_DATA = 2,
several
                                       # files with the extensions
"_1" to
                                       # "_R" will be created).

PRINT_EVERY_PERM 0                     # Print every tested
permutation of the
                                       # runs and the
corresponding value of
                                       # SSC to a file specified
by
                                       # EVERY_PERMFILE (0 = don't
print,
                                       # 1 = print).
                                       # Note that printing may
result in a
                                       # very large file.

EVERY_PERMFILE arabid.every_permfile   # Every tested permutation
of the runs
                                       # and the corresponding SSC
will be
                                       # printed here.

PRINT_RANDOM_INPUTORDER 0              # Print random input orders
of runs to
                                       # RANDOM_INPUTORDER (0 =
don't print,
                                            # 1 = print). This
option is only
```

```
                                                # available if
GREEDY_OPTION = 2.

RANDOM_INPUTORDERFILE arabid.random_inputorderfile # Every random
input order
                                                # of the runs
(generated by CLUMPP if
                                                # GREEDY_OPTION = 2)
will be printed
                                                # here.


# --------------- Advanced options ------------------------------
--------------

OVERRIDE_WARNINGS 0                     # This option allows the
user to
                                        # override non-crucial
warnings from
                                        # the program (0 allow
warnings, 1 do
                                        # not issue non-crucial
warnings).

ORDER_BY_RUN 1                          # Permute the clusters of
the output
                                        # files by the specified
run. (0 to
                                        # not specify a run, 1 to R
specifies
                                        # a run in the INDFILE or
POPFILE).


# --------------- Additional comments --------------------------
--------------

# The term ''permutation'' is used in two different contexts,
permutations of
# membership coefficients, or clusters, and permutations of runs.

# For example, if the datafile has has data A B C D E (each
letter indicates a
# column corresponding to a cluster), then permutation 3 2 5 1 4
of the
# clusters means C B E A D.

# Permutation 4 1 2 3 of runs 1-4 would mean start with run 4,
then run 1, then
```

```
# run 2, and then run 3.


# --------------- Command line arguments -----------------------
--------------

# -i INDFILE
# -p POPFILE
# -o OUTFILE
# -j MISCFILE
# -k K
# -c C
# -r R
# -m M
# -w W
# -s S
# ---------------------------------------------------------------
--------------
```

DISTRUCT PARAMETER FILE K=3

```
PARAMETERS FOR THE PROGRAM distruct.  YOU WILL NEED TO SET THESE
IN ORDER TO RUN THE PROGRAM.

"(int)" means that this takes an integer value.
"(B)"   means that this variable is Boolean
        (1 for True, and 0 for False)
"(str)" means that this is a string (but not enclosed in quotes)
"(d)"   means that this is a double (a real number).

Data settings

#define INFILE_POPQ        k3.popq      // (str) input file of
population q's
#define INFILE_INDIVQ      k3.indivq    // (str) input file of
individual q's
#define INFILE_LABEL_BELOW k3.watersheds  // (str) input file of
labels for below figure
#define INFILE_LABEL_ATOP  k3.regions      // (str) input file
of labels for atop figure
#define INFILE_CLUST_PERM  k3.perm     // (str) input file of
permutation of clusters to print
#define OUTFILE            k3.ps       //(str) name of output
file

#define K  3    // (int) number of clusters
```

```
#define NUMPOPS 3    // (int) number of pre-defined populations
#define NUMINDS 146  // (int) number of individuals

Main usage options

#define PRINT_INDIVS      1  // (B) 1 if indiv q's are to be
printed, 0 if only population q's
#define PRINT_LABEL_ATOP  0  // (B) print labels above figure
#define PRINT_LABEL_BELOW 1  // (B) print labels below figure
#define PRINT_SEP         1  // (B) print lines to separate
populations

Figure appearance

#define FONTHEIGHT 6  // (d) size of font
#define DIST_ABOVE 5  // (d) distance above plot to place text
#define DIST_BELOW -7 // (d) distance below plot to place text
#define BOXHEIGHT  36 // (d) height of the figure
#define INDIVWIDTH 1.5     // (d) width of an individual


Extra options

#define ORIENTATION 0      // (int) 0 for horizontal orientation
(default)
                     //       1 for vertical orientation
                     //          2 for reverse horizontal
orientation
                            //       3 for reverse vertical
orientation
#define XORIGIN 75         // (d) lower-left x-coordinate of
figure
#define YORIGIN 250        // (d) lower-left y-coordinate of
figure
#define XSCALE 1      // (d) scale for x direction
#define YSCALE 1      // (d) scale for y direction
#define ANGLE_LABEL_ATOP 60// (d) angle for labels atop figure
(in [0,180])
#define ANGLE_LABEL_BELOW 60    // (d) angle for labels below
figure (in [0,180])
#define LINEWIDTH_RIM  3   // (d) width of "pen" for rim of box
#define LINEWIDTH_SEP 0.3  // (d) width of "pen" for separators
between pops and for tics
#define LINEWIDTH_IND 0.3  // (d) width of "pen" used for
individuals
#define GRAYSCALE 0           // (B) use grayscale instead of
colors
#define ECHO_DATA 1             // (B) print some of the data to
the screen
```

```
#define REPRINT_DATA 1          // (B) print the data as a
comment in the ps file
#define PRINT_INFILE_NAME 0     // (B) print the name of
INFILE_POPQ above the figure
                                //     this option is meant for
use only with ORIENTATION=0
#define PRINT_COLOR_BREWER 1    // (B) print ColorBrewer settings
in the output file
                                //     this option adds 1689
lines and 104656 bytes to the output
                                //     and is required if using
ColorBrewer colors


Command line options:

-d drawparams
-K K
-M NUMPOPS
-N NUMINDS
-p input file (population q's)
-i input file (individual q's)
-a input file (labels atop figure)
-b input file (labels below figure)
-c input file (cluster permutation)
-o output file
```

DISTRUCT PARAMETER FILE K=2

```
PARAMETERS FOR THE PROGRAM distruct.  YOU WILL NEED TO SET THESE
IN ORDER TO RUN THE PROGRAM.

"(int)" means that this takes an integer value.
"(B)"   means that this variable is Boolean
        (1 for True, and 0 for False)
"(str)" means that this is a string (but not enclosed in quotes)
"(d)"   means that this is a double (a real number).

Data settings

#define INFILE_POPQ        k2.popq      // (str) input file of
population q's
```

```
#define INFILE_INDIVQ      k2.indivq    // (str) input file of
individual q's
#define INFILE_LABEL_BELOW k2.names  // (str) input file of
labels for below figure
#define INFILE_LABEL_ATOP  k2.regions      // (str) input file
of labels for atop figure
#define INFILE_CLUST_PERM  k2.perm     // (str) input file of
permutation of clusters to print
#define OUTFILE            k2.ps       //(str) name of output
file

#define K  2    // (int) number of clusters
#define NUMPOPS 13    // (int) number of pre-defined populations
#define NUMINDS 123  // (int) number of individuals

Main usage options

#define PRINT_INDIVS      1  // (B) 1 if indiv q's are to be
printed, 0 if only population q's
#define PRINT_LABEL_ATOP  0  // (B) print labels above figure
#define PRINT_LABEL_BELOW 1  // (B) print labels below figure
#define PRINT_SEP         1  // (B) print lines to separate
populations

Figure appearance

#define FONTHEIGHT 6  // (d) size of font
#define DIST_ABOVE 5  // (d) distance above plot to place text
#define DIST_BELOW -7 // (d) distance below plot to place text
#define BOXHEIGHT  36 // (d) height of the figure
#define INDIVWIDTH 1.5     // (d) width of an individual


Extra options

#define ORIENTATION 0      // (int) 0 for horizontal orientation
(default)
                     //       1 for vertical orientation
                     //          2 for reverse horizontal
orientation
                            //       3 for reverse vertical
orientation
#define XORIGIN 75         // (d) lower-left x-coordinate of
figure
#define YORIGIN 250        // (d) lower-left y-coordinate of
figure
#define XSCALE 1      // (d) scale for x direction
#define YSCALE 1      // (d) scale for y direction
```

```
#define ANGLE_LABEL_ATOP 60// (d) angle for labels atop figure
(in [0,180])
#define ANGLE_LABEL_BELOW 60    // (d) angle for labels below
figure (in [0,180])
#define LINEWIDTH_RIM  3    // (d) width of "pen" for rim of box
#define LINEWIDTH_SEP 0.3   // (d) width of "pen" for separators
between pops and for tics
#define LINEWIDTH_IND 0.3   // (d) width of "pen" used for
individuals
#define GRAYSCALE 0           // (B) use grayscale instead of
colors
#define ECHO_DATA 1             // (B) print some of the data to
the screen
#define REPRINT_DATA 1          // (B) print the data as a
comment in the ps file
#define PRINT_INFILE_NAME 0     // (B) print the name of
INFILE_POPQ above the figure
                                //     this option is meant for
use only with ORIENTATION=0
#define PRINT_COLOR_BREWER 1    // (B) print ColorBrewer settings
in the output file
                                //     this option adds 1689
lines and 104656 bytes to the output
                                //     and is required if using
ColorBrewer colors


Command line options:

-d drawparams
-K K
-M NUMPOPS
-N NUMINDS
-p input file (population q's)
-i input file (individual q's)
-a input file (labels atop figure)
-b input file (labels below figure)
-c input file (cluster permutation)
-o output file
```

REFERENCES

Aguilar, A., & Jones, W. J. (2009). Nuclear and mitochondrial diversification in two native California minnows: insights into taxonomic identity and regional phylogeography. *Molecular Phylogenetics and Evolution, 51*(2), 373-381.

Al-Rabab'ah, M., & Williams, C. (2002). Population dynamics of Pinus taeda L. based on nuclear microsatellites. *Forest Ecology and Management, 163*(1-3), 263-271.

Ardren, W., Borer, S., Thrower, F., Joyce, J., & Kapuscinski, A. (1999). Inheritance in 12 microsatellite loci in Oncorhynchus mykiss. *Journal of Heredity, 90*, 529-536.

Avise, J. (1991). Ten unorthodox perspectives on evolution prompted by comparative population genetic findings on mitochondrial DNA. *Annual Review of Genetics, 25*, 45-69.

Avise, J. C. (1994). *Molecular Markers, Natural History and Evolution.* New York: Chapman and Hall.

Avise, J. C. (2000). *Phylogeography: the history and formation of species.* Cambridge, MA: Harvard University Press.

Avise, J. C. (2009). Phylogeography: retrospect and prospect. *Journal of Biogeography, 36*, 3-15.

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., . . . Saunders, N. C. (1987). Intraspecific Phylogeography: The Mitochondrial DNA Bridge Between Population Genetics and Systematics. *Annual Review of Ecology and Systematics, 18*, 489-522.

Avise, J. C., Ball, R. M., & Arnold, J. (1988). Current versus historical population sizes in vertebrate species with high gene flow: a comparison based on mitochondrial DNA lineages and inbreeding theory for neutral mutations. *Molecular Biology Evolution, 5*(4), 331-344.

Avise, J. C., Neigel, J. E., & Arnold, J. (1984). Demographic influences on mitochondrial DNA lineage survivorship in animal populations. *Journal of Molecular Evolution, 20*(2), 99-105.

Balloux, F., & Goudet, J. (2002, February). The estimation of population differentiation with microsatellite markers. *Molecular Ecology, 11*(2), 155-165.

Bensasson, D., Zhang, D., Hart, l. D., & Hewitt, G. (2001). Mitochondrial pseudogenes: Evolution's misplaced witnesses. *Trends in Ecology and Evolution, 16*, 314-321.

Bensasson, D., Zhang, D.-X., & Hewitt, G. M. (2000). Frequent Assimilation of Mitochondrial DNA by Grasshopper. *Molecular Biology and Evolution, 17*(3), 406-415.

Borer, S., Miller, L., & Kapuscinski, A. (1999). Microsatellites in walleye Stizostedion vitreum. *Molecular Ecology, 8*(2), 336-338.

Brodie III, E. D., & Brodie Jr., E. D. (1990). Tetrodotoxin Resistance in Garter Snakes: An Evolutionary Response of Predators to Dangerous Prey. *Evolution, 44*(3), 651-659.

Brown, J., Leebens-Mack, J., Thompson, J., Pellmyr, O., & Harrison, R. (1997). Phylogeography and host association in a pollinating seed parasite Greya politella (Lepidoptera Prodoxidae). *Molecular Ecology, 6*, 215-224.

Brown, W. M., Matthew George, J., & Wilson, A. C. (1979). Rapid evolution of animal mitochondrial DNA. *Procedings of the National Academy of Sciences, 76*(4), 1967-1971.

Burns, K. J., & Barhoum, D. N. (2006). Population-level history of the wrentit (Chamaea fasciata): Implications for comparative phylogeography in the California Floristic Province. *Molecular Phylogenetics and Evolution, 38*(1), 117-129.

Chatzimanolis, S., & Caterino, M. S. (2007). Toward a better understanding of the "Transevese Range Break": Lineage Diversification in Southern California. *Evolution, 61*(9), 2127-2141.

Colburn, I. (2006). The Role of Antecedent Rivers in Shaping the Orange/Los Angeles Coastal Plain. *Society for Sedimentary Geology*.

Cornelius, R. (1969). *The systematics and zoogeography of Rhinichthys osculus (Girard) in Southern California.* Fullerton: California State College.

Dawid, I., & Blackler, A. (1972). Maternal and cytoplasmic inheritance of mitochondrial DNA in Xenopus. *Developmental Biology, 2*, 152-161.

Di Rienzo, A., Peterson, A. C., Garza, J. C., Valdes, A. M., Slatkin, M., & Freimer, N. B. (1994). Mutational processes of simple-sequence repeat loci in human populations. *Proceedings of the National Academy of Sciences, 91*(8), 3166-3170.

Dupanloup, I., Schneider, S., & Excoffier, L. (2002). A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology, 11*(12), 2571-2581.

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetic Resources, 4*(2), 359-361.

Epifanio, J., Johnson, J., & Kassler, T. (2003). *Genetic Analysis of Tennessee Sport Fisheries (2000 -2003):Largemouth Bass, Smallmouth Bass, Rainbow Trout, &Brown Trout.* University of Illinois atUrbana-Champaign. Nashville: Tennessee Wildlife Resources AgencyEllington Agricultural Center.

Estoup, A., & Angers, B. (1998). Microsatellites and minisatellites for molecular ecology: theoretical and empirical considerations. *Advances in Molecular Ecology*, 55-86.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology, 14*(8), 2611–2620.

Excoffier, L., & Slatkin, M. (1995). Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Molecular Biology Evolution, 12*(5), 921-927.

Excoffier, L., Laval, G., & Schneider, S. (2005). Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics, 1*, 27-50.

Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of Molecular Variance Inferred from Metric Distances among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data. *Genetics, 131*(2), 479-491.

Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. *Genetics, 164*(4), 1567-1587.

Falush, D., Stephens, M., & Pritchard, J. K. (2007). Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Resources, 7*(4), 574-578.

Fernandes-Matioli, F., Matioli, S., & Almeida-Toledo, L. (2000). Species diversity and geographic distribution of Gymnotus (Pisces: Gymnotiformes) by nuclear (GGAC)n microsatellite analysis. *Genetics and Molecular Biology, 23*, 803-807.

Forbes, S. H., & Boyd, D. K. (1996). Genetic variation of naturally colonizing wolves in the Central Rocky Mountains. *Conservation Biology, 10*, 1082-1090.

Frissell, C. A., Liss, W. J., Warren, C. E., & Hurley, M. D. (1986). A hierarchical framework for stream habitat classification: Viewing streams in a watershed context. *Environmental Management, 10*(2), 199-214.

Garcia-Moreno, J., Matocq, M. D., Roy, M. S., Geffen, E., & Wayne, R. K. (1996). Relationships and Genetic Purity of the Endangered Mexican Wolf Based on Analysis of Microsatellite Loci. *Conservation Biology, 10*(2), 376-389.

Girard, P., & Angers, B. (2006). Characterization of microsatellite loci in longnose dace (Rhinichthys cataractae) and interspecific amplification in five other Leuciscinae species. *Molecular Ecology Notes, 6*(1), 69-71.

Hardy, O. J., Charbonnel, N., Fréville, H., & Heuertz, M. (2003). Microsatellite Allele Sizes: A Simple Test to Assess Their Significance on Genetic Differentiation. *Genetics, 163*(4), 1467-1482.

Harrison, R. (1989). Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. *Trends in Ecology and Evolution, 4*, 6-11.

Harrison, S. (1999). Local and Regional Diversity in a Patchy Landscape: Native, Alien, and Endemic Herbs on Serpentine. *Ecology, 80*(1), 70-80.

Hartl, D. L., & Clark, A. G. (2006). *Principles of Population Genetics* (4 ed.). Sunderland, MA: Sinauer Associates, Inc.

Hedrick, P. W. (1998). *Genetics of Populations* (3 ed.). Sudbury, Massachusetts: JONES AND BARTLETT PUBLISHERS.

Hoekzema, K., & Sidlauskas, B. L. (2014). Molecular phylogenetics and microsatellite analysis reveal cryptic species of speckled dace (Cyprinidae: Rhinichthys osculus) in Oregon's Great Basin. *Molecular Phylogenetics and Evolution, 77*, 238-250.

Hubbs, C., Miller, R., & Hubbs, L. (1974). Hydrographic History and Relict Fishes of the North-Central Great Basin. *California Academy of Science, 7*, 1-254.

Hubisz, M. J., Falush, D., Stephens, M., & Pritchard, J. K. (2009). Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources, 9*(5), 1322–1332.

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics, 23*(14), 1801-1806.

Jarne, P., & Lagoda, P. J. (1996). Microsatellites, form molecules to populations and back. *Trends in Evolution and Ecology, 11*(10), 424-429.

Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics, 27*(21), 3070-3071.

Jordan, D., Evermann, B., & Clark, H. (1930). Check list of the fishes and fishlike vertebrates of North and Middle America north of the northern boundary of

Venezuela and Columbia. *Report of the United States Commisioner of Fisheries for the fiscal year 1928*, 670.

Kinziger, A. P., Nakamoto, R. J., Anderson, E. C., & Harvey, B. C. (2011). Small founding number and low genetic diversity in an introduced species exhibiting limited invasion success (speckled dace, Rhinichthys osculus). *Ecology and Evolution, 1*(1), 73-84.

Kumar, S., & Subramanian, S. (2001). Mutation rates in mammalian genomes. *Procedings of the National Academy of Sciences*, 803-808.

Lee, D., Gilbert, C., Hocutt, C., Jenkins, R., McAllister, D., & Stauffer, J. J. (1980). *Atlas of North American freshwater fishes.* Raleigh, NC: North Carolina State Museum of Natural History.

Maldonado, J. E., Davila, F. O., Stewart, B. S., Geffen, E., & Wayne, R. K. (1995). Intraspecific Genetic Differntiation in California Sea Lions (Zalophus californianus) from Southern California and the Gulf of California. *Marine Mammal Science, 11*(1), 46-58.

McConnell, S., O'Reilly, P., Hamilton, L., Wright, J., & Bentzen, P. (1995). Polymorphic microsatellite loci from Atlantic salmon (Salmo salar): genetic differentiation of North American and European populations. *Canadian Journal of Fisheries and Aquatic Sciences, 52*(9), 1863-1872.

Metcalf, A. E., Nunney, L., & Hyman, B. C. (2001). Geographic Patterns of Genetic Differentiation within the Restricted Range of the Endangered

Stephens' Kangaroo Rat Dipodomys stephensi. *Evolution, 55*(6), 1233-1244.

Meyes, N., Mittermeier, R. A., Mittermeier, C. G., Fonseca, G. A., & Kent, J. (2000). Biodiversity Hotspots for Conservation Priorities. *Nature, 403*, 853-858.

Mila, B., Girman, D. J., Kimura, M., & Smith, T. B. (2000). Genetic evidence for the effect of a postglacial population expansion on the phylogeography of a North American songbird. *The Proceeding of the Royal Society of London, 267*, 1033-1040.

Miller, R. R. (1958). Origin and affinities of the freshwater fish fauna of Western North America. (C. Hubbs, Ed.) *Zoogeography*(51), 187-222.

Miller, R. R., & Miller, R. G. (1948). The Contribution of the Columbia River System to the Fish Fauna of Nevada: Five Species Unrecorded from the State. *Copeia, 1948*(3), 174-187.

Moritz, C., Dowling, T. E., & Brown, W. M. (1987). Evolution of Animal Mitochondrial DNA: Relevance for Population Biology and Systematics. *Annual Review of Ecology and Systematics, 18*, 269-292.

Moyle, P. B., & Marchetti, M. P. (2006). Predicting Invasion Success Freshwater Fishes in California as a Model. *Bioscience, 56*(6), 515-524.

Moyle, P. B., Yoshiyama, R. M., Williams, J. E., & Wikramanayake, E. D. (1995). *FISH SPECIES OF SPECIAL CONCERN IN CALIFORNIA.* Rancho Cordova: California Department of Fish and Game.

Nei, M. (1987). *Molecular Evolutionary Genetics.* New York: Columbia University Press.

Nei, M., Maruyama, T., & Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. *Evolution, 29*(1), 1-10.

Neigel, J. E. (1997). A Comparison of Alternative Strategies for Estimating Gene Flow from Genetic Markers. *Annual Review of Ecology and Systematics, 28*, pp. 105-128.

Nico, L., & Fuller, P. (2015). *Rhinichthys osculus*. Retrieved from USGS Nonindigenous Aquatic Species Database: http://nas.er.usgs.gov/

Nunziata, S. O., Lance, S. L., Jones, K. L., Nerkowski, S. A., & Metcalf, A. E. (2013). Developement and characterization of 23 microsatellite markers for Rhinichthys osculus using paired-end Illumina shotgun sequencing. *Conservation Biology Notes*.

Oakey, D. D., Douglas, M. E., & Douglas, M. R. (2004). Small Fish in a Large Landscape: Diversification of Rhinichthys osculus (Cyprinidae) in Western North America. *Copeia, 2001*(2), 207-221.

O'Connell, M., & Wright, J. M. (1997). Microsatellite DNA in fishes. *Reviews in Fish Biology and Fisheries, 7*(3), 331-363.

O'Reilly, P., & Wright, J. M. (1995). The evolving technology of DNA fingerprinting and its application to fisheries and aquaculture. *Jounral of Fish Biology, 47*, 29-55.

Paetkau, D., Calvert, W., Stiriling, I., & Strobeck, C. (1995). Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology, 4*(3), 347-354.

Peakall, R., & Smouse, P. E. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes, 6*(1), 288-295.

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics, 28*(19), 2537-2539.

Pfrender, M. E., Hicks, J., & Lynch, M. (2004). Biogeographic patterns and current distribution of molecular-genetic variation among populations of speckled dace, Rhinichthys osculus (Girard). *Molecular Phylogenetics and Evolution, 30*(3), 490-502.

Phillipsen, I. C., & Metcalf, A. E. (2009). Phylogeography of a stream-dwelling frog (Pseudacris cadaverina) in Southern California. *Molecular Phylogenetics and Evolution, 53*, 152-170.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics, 155*(2), 945-959.

Rodriguez-Robles, J. A., Denardo, D. F., & Staub, R. E. (1999). Phylogeography of the California mountain kingsnake,Lampropeltis zonata (Colubridae). *Molecular Ecology, 8*, 1923-1934.

Rosenberg, N. A. (2004). DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Resources, 4*(1), 137-138.

Santa Ana River Watershed Project Authority. (2004). *Old, Grand Prix, and Padua Fires (October, 2003): Burn Impacts to Water Systems and Resources. Santa Ana River Watershed Area. San Bernardino National Forest, California. Summary Report.* Santa Ana Watershed Project Authority.

Santa Ana Watershed Association. (2009). *Santa Ana Watershed Association 2008-2009 Annual Report.* Chino: Santa Ana Watershed Association.

Santos, N. R., Katz, J. V., Moyle, P. B., & Viers, J. H. (2014). *RHINICHTHYS OSCULUS SUBSPECIES.* Retrieved from UC DAVIS PISCES: http://pisces.ucdavis.edu/content/rhinichthys-osculus-subspecies-2

Schlötterer, C., & Harr, B. (2000). Drosophila virilis has long and highly polymorphic microsatellites. *Molecular Biology and Evolution, 17*(11), 1641-1646.

Scott, W., & Crossman, E. (1973). Freshwater Fishes of Canada. *Fisheries Research Board of Canada Bulletin, 184*, 966.

Shannon1. (2015, March 27). *Map of the Santa Ana River basin.* Retrieved from Newkis: http://www.newikis.com/en/commons/File:Santa-ana-river-new.jpg

Shi, W., Kerdelhue, C., & Ye, H. (2012). Genetic Structure and Inferences on Potential Source Areas for Bactrocera dorsalis (Hendel) Based on Mitochondrial and Microsatellite Markers. *PLoS One, 7*(5).

Simmons, A. M., Berendzen, P. B., & Mayden, R. L. (2003). Molecular systematics of North American phoxinin genera (Actinopterygii: Cyprinidae) inferred from mitochondrial 12S and 16S ribosomal RNA sequences. *Zoological Journal of the Linnean Society, 139*, 63-80.

Simon, C. (1991). Molecular systematics at the species boundary: exploiting conserved and variable regions of the mitochondrial genome of animals via direct sequencing from amplified DNA. *Molecular Techniques in Taxonomy*, 33-71.

Sites Jr., W. J., & Crandall, K. A. (1997). Testing Species Boundaries in Biodiversity Studies. *Conservation Biology, 11*(6), 1289-1297.

Smith, G. R., & Dowling, T. E. (2008). Correlating hydrographic events and divergence times of speckled dace (Rhinichthys: Teleostei: Cyprinidae) in the Colorado River drainage. *The Geological Society of America - Special Paper*, 301-315.

Spinks, P. Q., Thomson, R. C., & Shaffer, H. B. (2010). Nuclear gene phylogeography reveals the historical legacy of an ancient inland sea on lineages of the western pond turtle, Emys marmorata in California. *Molecular Ecology, 19*(3), 542-556.

Tan, A.-M., & Wake, D. B. (1995). MtDNA phylogeography of the California newt, Taricha torosa (Caudata, Salamandridae). *Molecular Phylogenetics and Evolution, 4*(4), 383-394.

Turner, T. F., Dowling, T. E., Broughton, R. E., & Gold, J. R. (2004). Variable microsatellite markers amplify across divergent lineages of cyprinid fishes (subfamily Leusicinae). *Conservation Genetics, 5*(2), 279-281.

Vandergast, A. G., Bohonak, A. J., Weissman, D. B., & Fisher, R. N. (2007). Understanding the genetic effects of recent habitat fragmentation in the context of evolutionary history: phylogeography and landscape genetics of a southern California endemic Jerusalem cricket (Orthoptera: Stenopelmatidae: Stenopelmatus). *Molecular Ecology, 16*, 977-992.

Vila, C., Amorim, I. R., Leonard, J., Posada, D., Castroviejo, J., Petrucci-Fonseca, F., . . . Wayne, R. (1999). Mitochondrial DNA phylogeography and population history of the grey wolf. *Molecular Ecology, 8*, 2089-2103.

Ward, R., Bowers, K., Hensley, R., Mobley, B., & Belouski, E. (2007). Genetic variability in spotted seatrout (Cynoscion nebulosus), determined with microsatellite DNA markers. *Fishery Bulletin, 105*, 197-206.

Weir, B., & Cockerham, C. C. (1984). Estimating F-Statistics for the Analysis of Population Structure. *Evolution, 38*(6), 1358-1370.

Wright, S. (1931). Evolution in Mendelian populations. *Genetics, 16*, 97-159.

Zhang, D.-X., & Hewitt, G. M. (1996). Nuclear integrations: challenges for mitochondrial DNA markers. *Trends in Ecology and Evolution, 11*(6), 247-251.

Zink, R. M. (1996). Comparative Phylogeography in North American Birds. *Evolution, 50*(1), 308-317.